A T cell receptor locus encodes a malaria-specific immune response gene

Natalija Van Braeckel-Budimir[1]*, Stephanie Gras[2,3]*, Kristin Ladell[4]*, Tracy M. Josephs[2,3], Lecia Pewe[1], Stina L. Urban[1], Kelly L. Miners[4], Carine Farenc[3], David A. Price[4,5+], Jamie Rossjohn[2,3,4+], John T. Harty[1,6,7+]

Running title: An *Ir* gene in a TCR locus

[1]Department of Microbiology, University of Iowa, Iowa City, IA 52242, USA;
[2]Australian Research Council Centre of Excellence for Advanced Molecular Imaging, Monash University, Clayton, Victoria 3800, Australia;
[3] Infection and Immunity Program and Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia;
[4]Division of Infection and Immunity, Cardiff University School of Medicine, Cardiff CF14 4XN, UK;
[5]Human Immunology Section, Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA;
[6] Department of Pathology, University of Iowa, Iowa City, IA 52242, USA;
[7]Interdisciplinary Program in Immunology, University of Iowa, Iowa City, IA 52242, USA.

*These authors contributed equally.
[+]Senior authors.

Correspondence: JTH (john-harty@uiowa.edu), JR (jamie.rossjohn@monash.edu) or DAP (priced6@cardiff.ac.uk).

Lead contact: JTH (john-harty@uiowa.edu).

SUMMARY

Immune response (*Ir*) genes, originally proposed by Baruj Benacerraf to explain differential antigen-specific responses in animal models, have become synonymous with the major histocompatibility complex (MHC). We discovered a non-MHC-linked *Ir* gene in a T cell receptor (TCR) locus that was required for $CD8^+$ T cell responses to the *Plasmodium berghei* $GAP50_{40-48}$ epitope in mice expressing the MHC class I allele H-$2D^b$. $GAP50_{40-48}$-specific $CD8^+$ T cell responses emerged from a very large pool of naive $V\beta8.1^+$ precursors, which dictated susceptibility to cerebral malaria and conferred protection against recombinant *Listeria monocytogenes* infection. Structural analysis of a prototypical $V\beta8.1^+$ TCR-H-$2D^b$-$GAP50_{40-48}$ ternary complex revealed that germline-encoded complementarity-determining region 1β residues present exclusively in the Vβ8.1 segment mediated essential interactions with the $GAP50_{40-48}$ peptide. Collectively, these findings demonstrated that *Vβ8.1* functioned as an *Ir* gene that was indispensable for immune reactivity against the malaria $GAP50_{40-48}$ epitope.

KEY WORDS

Immune response genes; TCR bias; naive $CD8^+$ T cells; malaria; experimental cerebral malaria.

INTRODUCTION

Genetic regulation of the immune system is a central determinant of health and disease. Based on the observation that random-bred guinea pigs challenged with simple antigens segregate into responder and non-responder groups, Benacerraf and colleagues proposed the notion of autosomal dominant immune response (*Ir*) genes (Benacerraf and Germain, 1978). *Ir* genes were first identified as major histocompatibility complex (MHC)-linked (McDevitt and Chinitz, 1969), and subsequently found to encode specific allotypes of antigen-presenting molecules, such as MHC class II (MHC II) (Benacerraf, 1974; Benacerraf and McDevitt, 1972). Although unidentified at the time, a determinative role was also suggested for the putative T cell receptor (TCR) (Benacerraf and Germain, 1978). MHC II-mediated antigen presentation is required to activate T helper cells and generate antibody responses to T-dependent antigens (Owens and Zeine, 1989). In addition, the host must assemble TCRs capable of engaging specific MHC-peptide (MHCp) complexes with sufficient avidity to trigger immune reactivity (Davis et al., 2003; van der Merwe and Dushek, 2011). Numerous studies of inbred animals have linked the absence of specific immune responses with a lack of appropriate MHC alleles (Marshak et al., 1977; Zinkernagel, 1978). In contrast, despite major advances in our understanding of antigen recognition over the last four decades, it remains unclear if germline-encoded segments of the TCR can function as *Ir* genes.

$CD8^+$ T cells recognize MHC I-restricted peptides via heterodimeric TCRs (Davis and Bjorkman, 1988; Townsend et al., 1985). A vast number of different TCRs can be generated from a limited number of germline-encoded segments through the process of *V(D)J* gene recombination with junctional diversification and subsequent random pairing of the somatically rearranged TCRα and TCRβ chains (Cabaniols et al., 2001; Chothia et al., 1988; Davis and Bjorkman, 1988; Rossjohn et al., 2015; Turner et al., 2006). As a consequence, each individual harbors an extensive repertoire of naive $CD8^+$ T cells, which ensures broad recognition of a large number of foreign antigens presented by MHC I (Goldrath and Bevan, 1999). To meet the diversity criterion within space limitations, however, only a few naive precursors are specific for any given epitope (Blattman et al., 2002; Obar et al., 2008), and robust antigen-driven proliferation is

required to establish effector and memory $CD8^+$ T cell populations (Busch et al., 1998; Goldrath and Bevan, 1999). It is estimated that most naive antigen-specific repertoires in mice do not contain more than 10–300 $CD8^+$ T cells (Obar et al., 2008). Larger pre-immune repertoires comprising 1,000–1,500 naive $CD8^+$ T cells have been reported for the murine cytomegalovirus (MCMV) $M45_{985\text{-}993}$ and vaccinia virus (VacV) $B8R_{20\text{-}27}$ epitopes (Jenkins and Moon, 2012), but it remains unclear if these specific precursor pools define the upper limits of antigen reactivity in the post-thymic landscape of clonotypically distributed TCRs.

Clonal selection ensures the recruitment of biologically and structurally optimal immune receptors from the naive repertoire (Malherbe et al., 2004; Price et al., 2005), frequently leading to biased TCR usage among memory $CD8^+$ T cell populations (Miles et al., 2011; Turner et al., 2006). In extreme cases, non-peptidic antigens restricted by non-classical MHC molecules elicit innate-like responses dominated by semi-invariant TCRs (Bendelac et al., 1997; Godfrey et al., 2015; Van Rhijn et al., 2015). Here, we found that a similar phenomenon can regulate conventional $CD8^+$ T cell immunity. We demonstrated that an epitope from the *Plasmodium berghei* (*P. berghei*) ANKA glideosome-associated protein ($GAP50_{40\text{-}48}$), a key pathogenic target in experimental cerebral malaria (ECM), was associated with the largest naive $CD8^+$ T cell repertoire yet described in laboratory mice (Jenkins and Moon, 2012). We also revealed the molecular properties of these $GAP50_{40\text{-}48}$-specific TCRs, demonstrating exclusive use of the Vβ8.1 segment underpinned by direct interactions with the malarial peptide. Mice lacking the *Vβ8.1* gene did not respond to the $GAP50_{40\text{-}48}$ epitope after *P. berghei* infection and did not develop ECM. Moreover, the very large pool of naive precursors conferred enhanced control of primary infection with recombinant *Listeria monocytogenes* (*L. monocytogenes*) expressing the $GAP50_{40\text{-}48}$ epitope. Collectively, these findings extend the notion of *Ir* genes to incorporate germline-encoded components of antigen-specific TCRs.

RESULTS

GAP50$_{40-48}$-specific CD8$^+$ T cells exhibit an extreme TCR bias

ECM in susceptible C57Bl/6 (B6) mice infected with *P. berghei* ANKA (Engwerda et al., 2005) is a valuable model of severe malarial disease (Brewster et al., 1990). It is established that the development of ECM is critically dependent on pathogenic CD8$^+$ T cells (Amani et al., 2000; Haque et al., 2011; Yanez et al., 1996) expressing Vβ8$^+$ TCRs (Boubou et al., 1999; Mariotti-Ferrandiz et al., 2016), especially those specific for the H-2D^b-restricted GAP50$_{40-48}$ epitope (Howland et al., 2013). However, it is not known why GAP50$_{40-48}$-specific CD8$^+$ T cells are pathogenic in ECM. To address this issue, we set out to generate TCR retrogenic mice harboring monoclonal or oligoclonal CD8$^+$ T cell populations specific for individual epitopes derived from *P. berghei*, namely thrombospondin-related adhesion protein (TRAP)$_{130-138}$, sporozoite-specific protein 20 (S20)$_{318-326}$ and GAP50$_{40-48}$ (Holst et al., 2006; Hafalla et al., 2013; Howland et al., 2013). As a first step, we immunized three separate groups of B6 mice with peptide-pulsed dendritic cells (DCs) followed 7 days later by recombinant *L. monocytogenes* expressing the same epitope (LM-GAP50$_{40-48}$). This accelerated prime-boost approach (Badovinac et al., 2005) elicited large CD8$^+$ T cell responses specific for TRAP$_{130-138}$, S20$_{318-326}$ and GAP50$_{40-48}$ (Fig. S1A). We then used the corresponding MHC I tetramers and a panel of anti-mouse Vβ antibodies to profile the constituent malaria-specific TCRs.

CD8$^+$ T cells specific for TRAP$_{130-138}$ and S20$_{318-326}$ expressed several different TCR Vβ segments with distinct preferences (Fig. 1A and S1B). For example, almost 75% of the S20$_{318-326}$-specific repertoire was focused on Vβ8.1/8.2 and Vβ8.3, while the TRAP$_{130-138}$-specific repertoire was dominated by Vβ2 and Vβ7. In contrast, >99% of GAP50$_{40-48}$-specific CD8$^+$ T cells expressed Vβ8.1/8.2, which cannot be discriminated by antibody staining (Fig. 1A and S1B). These segments comprised <15% of the corresponding GAP50$_{40-48}$ tetramer$^-$ repertoires in the same mice (Fig. 1B). Moreover, an identical TCR bias was observed after inoculation of sporozoites by mosquito bite or immunization with radiation-attenuated sporozoites or DC-GAP50$_{40-48}$ alone (data not shown), and similar results were obtained with other strains of mice expressing the

MHC I allele H-2D^{b}, including CB6F1 and BALB.b (Fig. S1C). The GAP50$_{40\text{-}48}$ epitope therefore mobilized an almost exclusive repertoire of TCR Vβ-defined memory CD8^{+} T cells.

To gain a deeper understanding of this extreme bias, we sorted GAP50$_{40\text{-}48}$-specific CD8^{+} T cells directly *ex vivo* from DC-LM-GAP50$_{40\text{-}48}$-immunized B6 mice and performed an unbiased molecular analysis of all expressed *TRA* and *TRB* gene rearrangements (Quigley et al., 2011). Sequence analysis showed that >98% of transcripts encoded Vβ8.1 (Fig. 1C and Table S1). In contrast, a variety of *TRAV* gene segments were detected, indicating promiscuous pairing of diverse TCRα chains with a constrained repertoire of Vβ8.1^{+} TCRβ chains (Fig. 1D). Almost 40% of retrogenic T cells generated with a GAP50$_{40\text{-}48}$-specific TCRβ chain bound the GAP50$_{40\text{-}48}$ tetramer in naive mice, while only ~1% of retrogenic T cells generated with a TRAP$_{130\text{-}138}$-specific TCRβ chain bound the TRAP$_{130\text{-}138}$ tetramer (Fig. S2). These data suggested that the Vβ8.1 segment was a major recognition element for GAP50$_{40\text{-}48}$-specific CD8^{+} T cells.

Analysis of the third TCRβ chain complementarity-determining region (CDR3β) revealed additional features of the GAP50$_{40\text{-}48}$-specific repertoire (Fig. S3A–C, Table S1). In particular, we identified a common CDR3β length (~60% of translated sequences incorporated 14 amino acids) (Fig. S3A), an almost uniform bias towards *TRBD2* gene usage within a single reading frame (Fig. S3B) and a strong preference for *TRBJ* gene segments from the Jβ2 cluster (Fig. S3C). Moreover, the Vβ-Dβ junction lacked N additions, and a germline-encoded CDR3β motif (^{104}CASSDWG^{110}) was present in >95% of sequences (Table S1 and Fig. 1E). These data demonstrated that a highly convergent pattern of gene rearrangements “licensed” the Vβ8.1-driven immune response to the GAP50$_{40\text{-}48}$ epitope (Quigley et al., 2010).

The C-terminal region of the GAP50$_{40\text{-}48}$ peptide is a “hotspot” for TCR recognition

To understand the molecular basis of this extreme TCR bias, we first determined the structure of the H-2D^{b}-GAP50$_{40\text{-}48}$ complex at a resolution of 2.2 Å (Table S2). The bound peptide adopted a canonical extended conformation (Young et al., 1994), with four solvent-exposed amino acid residues (P4-L, where L is leucine; P6-A, where A is

alanine; P7-K, where K is lysine; and P8-Y, where Y is tyrosine) representing potential contacts for the TCR (Fig. 2A).

To determine which $GAP50_{40-48}$ residues were important for TCR recognition, we replaced individual amino acids with alanine (Ala), except for the wildtype (*wt*) Ala residue at position 6 (P6), which was replaced with serine (Ser) (Fig. 2B). We did not mutate the critical H-$2D^b$ anchor residues at P5 or P9 (Valkenburg et al., 2013). The *wt* and mutant peptides were then used at saturating concentrations in intracellular cytokine staining assays to determine the impact of each residue on the functional reactivity of $GAP50_{40-48}$-specific $CD8^+$ T cells (Fig. 2C and S4). Interferon-γ (IFNγ) production relative to the *wt* peptide was unaffected by the substitutions at P1, P2 and P3 (Fig. 2C). In contrast, the mutations at P4-L and P6-A diminished IFNγ production, while the mutations at P7-K and P8-Y abolished IFNγ production (Fig. 2C). The C-terminal P7-K and P8-Y residues were therefore crucial for $GAP50_{40-48}$-specific $CD8^+$ T cell activation and represented "hotspots" for TCR engagement (Fig. 2D).

$GAP50_{40-48}$ peptide "hotspots" interact exclusively with the TCRβ chain

Next, we determined the structure of a prototypical $V\beta8.1^+$ TCR, derived from the $GAP50_{40-48}$-specific $CD8^+$ T cell clone NB1 (Table S3), in complex with H-$2D^b$-$GAP50_{40-48}$ (Fig. 3A, Table S2). The NB1 TCR docked centrally atop the H-$2D^b$ cleft, with an angle of 43° and a buried surface area (BSA) of approximately 2,200 $Å^2$, values that fall within the range previously determined for TCR-MHCp complexes (Rossjohn et al., 2015) (Fig. 3B). All six CDR loops were involved to a varying extent in the interaction with H-$2D^b$-$GAP50_{40-48}$ (Fig. 3C). Namely, the CDR3α and CDR3β loops contributed 24% and 20% of the BSA, respectively, while the CDR1α and CDR2α loops each contributed ~17% of the BSA (Fig. 3C). Somewhat unexpectedly given the extreme TCR Vβ bias, the germline-encoded CDR1β and CDR2β loops each contributed only ~10% of the BSA. Thus, the cumulative β chain contribution to the overall BSA was substantially smaller (39%) relative to the cumulative α chain contribution (61%) (Fig. 3C). However, >95% of TCR-$GAP50_{40-48}$ peptide interactions were mediated by the TCRβ chain (Fig. 3D, Table S4), while the TCRα chain primarily contacted the MHC molecule (74% of the BSA) (Fig. 3E, Table S4).

The CDR1α loop stretched between the peptide-binding α-helices of H-2D^b, with tyrosine (Tyr) 28α playing a principal role by contacting glutamic acid (Glu)163, Glu166 and tryptophan (Trp)167 at the N-terminal end of the cleft (Fig. 3F, Table S4). The CDR2α loop sat above the α2-helix, with arginine (Arg)57α lying flat above residues Glu154 and Ala158 (Fig. 3G). The CDR3α loop contacted a large stretch of the α1-helix (spanning residues 58–72), with hydrophobic and hydrogen bond contacts exclusively mediated by germline-encoded residues from the Jα segment (^{108}YAQ^{110}, where Q is glutamine) (Fig. 3H). The NB1 TCRα chain engaged the H-2D^b molecule with a large footprint over the N-terminal region of the antigen-binding cleft (Fig. 3B). Mutational analyses of H-2D^b binding performed using three distinct GAP50$_{40\text{-}48}$-specific CD8$^+$ T cell clones further indicated that different Vβ8.1$^+$ TCRs used an identical MHC I docking strategy (Fig. S5, Table S5).

Interactions between the NB1 TCRα chain and the GAP50$_{40\text{-}48}$ peptide were limited to two van der Waals interactions between the CDR3α loop and P4-L (Table S4). In contrast, the NB1 TCR Vβ8.1 region dominated contacts with the bound epitope (44 of 46 contacts) (Table S4). Moreover, all three CDRβ loops contacted the C-terminus of the peptide (P6–8) (Fig. 3I–K), encompassing the previously identified “hotspots” for CD8$^+$ T cell recognition (Fig. 2D). The CDR1β loop interacted with both P7-K and P8-Y via its ^{28}N-DY^{31} motif (where N is asparagine and D is aspartic acid). More specifically, P7-K and P8-Y formed a notch in which aspartic acid (Asp)30β was inserted as a peg, forming a salt bridge with P7-K (Fig. 3J). In addition, asparagine (Asn)28β hydrogen bonded via its main chain with P7-K, while the aromatic group of Tyr31β sat above P8-Y (Fig. 3J). Contacts were also established between P8-Y and the CDR2β loop via Tyr57. Accordingly, P8-Y was closely sequestered upon TCR binding (Fig. 3J). The conserved CDR3β motif ^{108}DW^{109} (where W is tryptophan), which was most frequently germline-encoded within the *D* gene (Fig. 1E), interacted with P6-A and P7-K (Fig. 3K). The large Trp109β side chain lay flat on the top of P6-A, acting as a lid covering the central part of the epitope, while Asp108β and P7-K formed a salt bridge (Fig. 3K).

TCRβ chain interactions with the H-2D^b molecule were modest in comparison to the TCRα chain (Table S4). The CDR1β loop contacted only glutamine (Gln)72 via Tyr31β (Fig. 3I, J). The side chains of valine (Val)58β from the CDR2β loop abutted

Val76 from H-2D^b, establishing an interaction network extended by Tyr57β-mediated contacts with Gln72, Arg75 and Val76. The conserved CDR3β ^{108}DW^{109} motif from the *TRBD2* gene segment formed interactions with the hinge region of the α2-helix, whereby Trp109β nestled between histidine (His)155 and Ser150. Thus, germline-encoded TCRβ chain residues played a minor role in contacting the H-2D^b molecule.

The observation that the prototypical NB1 TCR interacted with GAP50$_{40\text{-}48}$ extensively via its β chain suggested that the peptide itself drove the extreme bias towards Vβ8.1. This notion contrasts with prevailing dogma, which asserts that germline bias arises as a consequence of allotype-specific MHC reactivity (Garcia et al., 2009).

CDR1 residues unique to Vβ8.1 are critical for GAP50$_{40\text{-}48}$ peptide recognition

These structural insights allowed us to identify candidate factors underlying the exclusive recruitment of Vβ8.1$^+$ TCR clonotypes into the immune repertoire. To determine the precise amino acid sequences involved in this process, we initially focused on the germline-encoded CDR1β and CDR2β loops. Sequence alignments revealed that the CDR1β loops encoded by six *TRBV* genes shared a common Asp30β, while the Asp30β-Tyr31β motif was unique to Vβ8.1. In the other five *TRBV* genes, Tyr31β was replaced by a smaller threonine (Thr)/Ser residue (Table S6A). Three *TRBV* genes shared a common Val58β residue in the encoded CDR2β loop, but only Vβ8.1 incorporated the additional Tyr57β required to sequester P8-Y (Table S6B).

Although the conserved CDR3β region interacted with both the peptide and H-2D^b, we wanted to determine if the germline-encoded Vβ8.1 residues were essential for TCR-mediated recognition of H-2D^b-GAP50$_{40\text{-}48}$. We therefore conducted affinity measurement studies using targeted mutants of the NB1 TCR. Ala mutagenesis was performed on residues Asp30β and Tyr31β in the CDR1β loop and residues Tyr57β and Val58β in the CDR2β loop (Fig. 4A and B). In addition, Tyr31β was mutated to Thr to mimic the CDR1β loop encoded by the five other *TRBV* genes that share Asp30β (Fig. 4A and B). Affinity values for each TCR mutant, measured by surface plasmon resonance (SPR), were compared with the affinity of the *wt* TCR. The results showed that mutations in the CDR2β loop exerted a moderate effect, decreasing binding affinity by 3-fold compared with the NB1 TCR (Fig. 4A and B). In contrast, mutation of Tyr31β

to either Ala or Thr in the CDR1β loop substantially reduced binding to H-2D^b-GAP50$_{40-48}$, resulting in Kd_{eq} values >200 μM (Fig. 4A and B). As shown in Fig. 3I and Table S4, Tyr31β contacted both H-2D^b and P8-Y. Additionally, mutation of Asp30β to Ala abolished binding to H-2D^b-GAP50$_{40-48}$ (Fig. 4A and B). The Asp30β residue contacted the peptide alone via a salt bridge with P7-K, one of the two defined "hotspot" residues for TCR recognition (Fig. 3J and Table S4).

The unique germline-encoded CDR1β ^{30}DY^{31} motif within the Vβ8.1 segment therefore underpinned both the extreme TCR bias and the residue-specific patterns of epitope recognition that characterized the CD8$^+$ T cell response to H-2D^b-GAP50$_{40-48}$.

The naive GAP50$_{40-48}$-specific CD8$^+$ T cell repertoire is extraordinarily large

Extreme TCR focusing described in the literature appears to result primarily from the selection of high-affinity clonotypes during repeated or persistent infection (Busch and Pamer, 1999; Malherbe et al., 2004; Price et al., 2005; Savage et al., 1999), but recombinatorial bias and structural constraints dictate the available pre-immune repertoire (Miles et al., 2011; Neller et al., 2015; Turner et al., 2006). We therefore adapted a tetramer-based enrichment protocol (Moon et al., 2007) to enumerate the naive repertoires specific for the *P. berghei* epitopes TRAP$_{130-138}$, S20$_{318-326}$ and GAP50$_{40-48}$. Parallel experiments were performed with tetramers representing epitopes derived from ovalbumin (H-2K^b-OVA$_{257-264}$) and lymphocytic choriomeningitis virus glycoprotein (H-2D^b-LCMV GP$_{33-41}$) to allow comparison with previously reported evaluations of the naive antigen-specific T cell repertoire (Jenkins and Moon, 2012).

Precursor numbers specific for OVA$_{257-264}$ and GP$_{33-41}$ were consistent with prior reports (183 ± 33 and 358 ± 40 cells, respectively, Fig. 5A) (Jenkins and Moon, 2012). The naive S20$_{318-326}$-specific repertoire was relatively small (~79 ± 19 cells), while fewer than 10 cells were counted in the TRAP$_{130-138}$-specific repertoire (Fig. 5A). In contrast, as described recently (Gordon et al., 2015), the GAP50$_{40-48}$-specific naive repertoire was extraordinarily large (2,935 ± 305 cells, Fig. 5A). The upper limit for a naive antigen-specific CD8$^+$ T cell repertoire in mice was previously established at 1,200–1,500 cells for the MCMV M45$_{985-993}$ and VacV B8R$_{20-27}$ epitopes (Jenkins and Moon, 2012). The GAP50$_{40-48}$-specific repertoire therefore constituted the largest naive

antigen-specific $CD8^+$ T cell pool yet described (Fig. 5B). Moreover, >99% of naive $GAP50_{40-48}$-specific $CD8^+$ T cells expressed Vβ8.1 (Fig. 5A). These findings indicated that the extreme TCR Vβ bias observed in the memory $CD8^+$ T cell pool did not evolve as a consequence of repertoire focusing in response to antigenic challenge, but instead reflected intrinsic recognition of the $GAP50_{40-48}$ epitope characterized by absolute dependence on the interaction with Vβ8.1.

Phenotypically, the $GAP50_{40-48}$-specific precursors resembled classical naive $CD8^+$ T cells, comprised of a large $CD44^{lo}$ population (>90%) and a small $CD44^{hi}$ "virtual memory" population (<10%) akin to OT-I and P14 TCR-Tg cells from naive mice (Fig. 5C) (Akue et al., 2012). Expression of the CD5 surface protein, which acts as a surrogate marker for the strength of TCR activation induced by self-derived MHCp complexes during thymic selection and frequently correlates with T cell avidity for antigen (Fulton et al., 2015; Mandl et al., 2013), was uniformly higher on naive $GAP50_{40-48}$-specific $CD8^+$ T cells compared with the majority of non-$GAP50_{40-48}$-specific naive $CD8^+$ T cells in the same host and mirrored levels expressed by OT-I cells, which express high-affinity TCRs (Fig. 5D) (Kedl et al., 2000). These data suggested that the extremely large pool of $V\beta8.1^+$ precursors specific for the $GAP50_{40-48}$ epitope arose as a consequence of strong thymic selection.

Absence of Vβ8.1 compromises the $GAP50_{40-48}$-specific $CD8^+$ T cell response

The results of our structural and mutagenesis studies indicated that immune responses to $GAP50_{40-48}$ required the presence of Vβ8.1. To test this hypothesis, we studied the generation of $GAP50_{40-48}$-specific $CD8^+$ T cell responses in C57/L ($H\text{-}2^b$) mice, which lack Vβ8 (Behlke et al., 1986). For this purpose, C57/L mice and control B6 mice were exposed to infection with 10^3 *P. berghei* ANKA sporozoites, a challenge that induces ECM in susceptible mouse strains.

In contrast to B6 mice, which generated $GAP50_{40-48}$-specific responses comprising 4–5% of the total $CD8^+$ T cell pool in peripheral blood and >10% of the total $CD8^+$ T cell pool in the brain by day 7 post-infection, the $GAP50_{40-48}$-specific response in C57/L mice was undetectable in both compartments (Fig. 6A and B). This finding strongly suggested that the $GAP50_{40-48}$-specific $CD8^+$ T cell response was critically

dependent on Vβ8.1. In addition, B6 x C57/L F1 mice mounted substantial Vβ8.1$^+$ GAP50$_{40\text{-}48}$-specific CD8$^+$ T cell responses after infection with LM-GAP50$_{40\text{-}48}$ or *P. berghei* ANKA (data not shown), excluding negative selection of GAP50$_{40\text{-}48}$-specific naive precursors in C57/L mice. However, C57/L mice lack other *Vβ* genes (*Vβ5*, *Vβ8*, *Vβ9*, *Vβ11* and *Vβ12*) (Behlke et al., 1986), which could potentially compromise the ability to mount alternative responses to GAP50$_{40\text{-}48}$. Countering this argument, the magnitude of the total *P. berghei*-specific CD8$^+$ T cell response, as detected by the frequency of CD8$^+$ T cells expressing surrogate activation markers (CD11a^{hi} CD8^{lo}) (Rai et al., 2009) at day 7 post-infection, was essentially identical in B6 and C57/L mice (Fig. 6C). Moreover, C57/L mice did not develop ECM, instead they survived beyond day 9 post-infection and succumbed to high-level parasitemia after ~3 weeks (Fig. 6D). Thus, the *Vβ8.1* gene was necessary for the generation of detectable GAP50$_{40\text{-}48}$-specific CD8$^+$ T cell responses and the development of ECM.

GAP50$_{40\text{-}48}$-specific CD8$^+$ T cells are pathogenic in ECM

A recent publication suggested that GAP50$_{40\text{-}48}$-specific CD8$^+$ T cells play a key role in the pathogenesis of ECM in susceptible B6 mice (Howland et al., 2013). Mechanistically, accumulation of GAP50$_{40\text{-}48}$-specific CD8$^+$ T cells in the brain could lead to eventual rupture of the blood-brain-barrier, which triggers well recognized neurological symptoms leading to rapid death (7–9 days post-infection) (Howland et al., 2013). In contrast to susceptible B6 mice, which all succumbed to ECM by day 7 post-infection with *P. berghei* ANKA-parasitized red blood cells (pRBCs), only 20% of CB6F1 mice succumbed to ECM (Fig. 6E). We therefore considered the possibility that a strain-specific difference in the size of the naive GAP50$_{40\text{-}48}$-specific CD8$^+$ T cell repertoire might underlie these differential outcomes. Tetramer pull-down of naive GAP50$_{40\text{-}48}$-specific CD8$^+$ T cells revealed that CB6F1 mice harbored ~3-fold fewer precursors than B6 mice (Fig. 6F). Moreover, this smaller naive pool resulted in a smaller GAP50$_{40\text{-}48}$-specific CD8$^+$ T cell response in CB6F1 mice compared with B6 mice at day 6 post-infection with *P. berghei* ANKA pRBCs (Fig. 6G). These results suggested that a GAP50$_{40\text{-}48}$-specific CD8$^+$ T cell response threshold might be required for the induction of ECM. To confirm a threshold-dependent pathogenic role for GAP50$_{40\text{-}48}$-specific

$CD8^+$ T cells in ECM, we generated a large $GAP50_{40-48}$-specific $CD8^+$ T cell response in CB6F1 or B6 mice prior to infection with *P. berghei* ANKA. We observed no meaningful difference in the onset of ECM symptoms or the mortality rate between previously immunized and non-immunized B6 mice, demonstrating that sufficient numbers of $GAP50_{40-48}$-specific effector $CD8^+$ T cells can be generated from the large pool of naive precursors in this susceptible strain (Fig. 6H). In contrast, prior immunization against $GAP50_{40-48}$ was required to render all CB6F1 mice susceptible to ECM after *P. berghei* ANKA infection (Fig. 6I). Collectively, these findings demonstrated that a numerical threshold delimited the pathogenicity of $GAP50_{40-48}$-specific $CD8^+$ T cells in the etiology of ECM.

$GAP50_{40-48}$-specific $CD8^+$ T cells protect against LM-$GAP50_{40-48}$

Although pathogenic in the context of ECM, we hypothesized that the rapid generation of a substantial $GAP50_{40-48}$-specific response from the large number of naive precursors may protect against infection controlled primarily by $CD8^+$ T cells. To address this possibility, we determined if B6 mice could better control primary infection with virulent LM-$GAP50_{40-48}$ compared with virulent *wt* LM.

To ensure comparability, we first established that the two different strains of LM exhibited similar virulence in the absence of a $GAP50_{40-48}$-specific $CD8^+$ T cell response. For this purpose, we infected BALB/c mice ($H\text{-}2^d$), which were non-responsive to the $H\text{-}2D^b$-restricted $GAP50_{40-48}$ epitope (Fig. 7A). Bacterial burden assessed in the spleens of infected mice 5 days after infection with 5×10^3 colony-forming units (CFU) of each strain was similar between the two groups, suggesting equivalent virulence *in vivo* (Fig. 7B). In contrast, we measured ~10-fold fewer bacteria in the spleens of LM-$GAP50_{40-48}$-infected B6 mice compared with *wt* LM-infected B6 mice (Fig. 7C–D). This superior control of bacterial infection in LM-$GAP50_{40-48}$-infected B6 mice was accompanied by a large $GAP50_{40-48}$ specific response, comprising >8% of the total $CD8^+$ T cell pool in the spleen (Fig. 7E). The expansion of $GAP50_{40-48}$-specific effector cells from a large pool of naive precursors therefore conferred protection in the context of an infection controlled by $CD8^+$ T cells.

DISCUSSION

Conceptual frameworks developed in the 1970s, prior to a detailed understanding of the molecular interactions that govern T cell antigen recognition, established the notion of genetically controlled immune responsiveness. Overwhelming experimental data have since accumulated to validate one original proposition that MHC-linked *Ir* genes dictate the presentation of specific antigens in immunogenic form (Benacerraf and Germain, 1978). There is also some evidence to support the idea that germline-encoded components of the TCR interact preferentially with defined MHC allotypes (Garcia et al., 2009; Scott-Browne et al., 2009), building on the earlier theoretical work of Niels Jerne (Jerne, 1971). However, it has not been shown previously that heritable elements of an antigen receptor can license immune reactivity, a scenario postulated almost four decades ago as an alternative model to explain the *Ir* gene phenomenon (Benacerraf and Germain, 1978). In this study, we demonstrated that $CD8^+$ T cell responses specific for the murine malaria epitope $GAP50_{40\text{-}48}$ were entirely dependent on amino acid residues unique to the TCR Vβ8.1 segment, thereby providing insights into the genetic mechanisms that control adaptive immunity.

As a consequence of structural constraints focused primarily on the CDR1β loop, both naive and antigen-expanded H-$2D^b$-restricted $GAP50_{40\text{-}48}$-specific $CD8^+$ T cells almost invariably express $V\beta8.1^+$ TCRs. Similarly extreme biases are typically associated with innate-like responses to non-peptidic antigens presented by non-classical MHC molecules, although immune reactivity among these unconventional T cell subsets is generally endowed by a conserved TCRα chain (Van Rhijn et al., 2015). Narrow TCR repertoires specific for peptide epitopes have previously been linked with low numbers of naive precursors (Moon et al., 2007) and clonal selection during the genesis of memory populations, especially in the presence of constant or repetitive antigen stimulation (Busch and Pamer, 1999; Price et al., 2005; Savage et al., 1999). Paradoxically, the pre-immune $GAP50_{40\text{-}48}$-specific repertoire was the largest yet described in mice, numbering ~3,000 cells in a naive B6 host. Moreover, this substantial precursor pool did not seem to arise as a consequence of expansion into the virtual memory subset (Haluszczak et al., 2009), as >90% of naive $GAP50_{40\text{-}48}$-specific $CD8^+$ T cells expressed low levels of CD44 (Akue et al., 2012). Although the self-derived

peptides associated with positive selection of these antigen-specific precursors remain unknown, the uniformly high expression of CD5 by naive $GAP50_{40-48}$-specific $CD8^+$ T cells suggested the existence of strong TCR-mediated interactions in the thymus (Mandl et al., 2013). It is also notable that *Vβ8* gene segments are highly represented in laboratory mouse strains (Wilson et al., 2001). High numbers of naive $GAP50_{40-48}$-specific $CD8^+$ T cells therefore most likely emerged as a consequence of both permissive thymic selection and the frequent generation of $V\beta8.1^+$ transcripts during the process of somatic recombination. In addition, the presence of a highly conserved Vβ-Dβ junction lacking N additions suggested that convergent gene rearrangements, which occur more commonly on a probabilistic basis (Quigley et al., 2010), effectively licensed the extreme penetrance of the *Vβ8.1* gene-encoded phenotype.

So what are the immunological consequences of such a large pool of antigen-specific precursors in the naive $CD8^+$ T cell compartment? Earlier work indicated that $CD8^+$ T cells expressing $V\beta8.1/V\beta8.2^+$ TCRs are highly prevalent during malaria infection and contribute to the pathogenesis of ECM (Mariotti-Ferrandiz et al., 2016). In addition, it had been shown previously that high-dose tolerization with the $GAP50_{40-48}$ peptide abrogates disease susceptibility in B6 mice (Howland et al., 2013). These observations can be explained by our finding that an extremely large pool of $V\beta8.1^+$ antigen-specific naive precursors (~3,000 cells) underpinned the immunodominant $CD8^+$ T cell response to $GAP50_{40-48}$ and the susceptibility of B6 mice to ECM. In contrast, CB6F1 mice harbored a diminished pool of antigen-specific naïve precursors (~1,000 cells), leading to smaller $GAP50_{40-48}$-specific $CD8^+$ T cell responses and relative resistance to ECM. This causal association was confirmed by the observation that increasing the number of $GAP50_{40-48}$-specific $CD8^+$ T cells in CB6F1 mice prior to infection rendered them susceptible to ECM. It remains unclear if similar germline biases can dictate immune responses to malarial antigens in humans, but our data nonetheless provide proof-of-concept that genetic associations with disease outcome can be extended to loci encoding components of the TCR. Moreover, the peak incidence of severe anemia during malaria infection occurs in children under 2 years of age, while cerebral malaria occurs most commonly in children aged 3–5 years who have prior exposure to infection (Struik and Riley, 2004). These epidemiological

patterns are consistent with the possibility that larger pre-infection repertoires of malarial antigen-specific T cells predispose to the development of cerebral malaria.

$GAP50_{40\text{-}48}$-specific memory $CD8^+$ T cells are not protective during the liver-stage of murine malaria (Doll et al., 2016; Horne-Debets et al., 2016; van der Heyde et al., 1993; Vinetz et al., 1990), and the capacity of $CD8^+$ T cells to protect during blood-stage malaria remains controversial (Horne-Debets et al., 2016; Imai et al., 2010; van der Heyde et al., 1993; Vinetz et al., 1990). Thus, the pathogenic effects of $GAP50_{40\text{-}48}$-specific $CD8^+$ T cells in ECM most likely resulted from a large population of precursors, which expanded rapidly after infection to generate substantial numbers of effector cells, leading to sustained immune activation without elimination of the pathogen. In contrast, rapid expansion from the same population of precursors enhanced immune control of primary infection with *L. monocytogenes* expressing the $GAP50_{40\text{-}48}$ epitope. An unusually large naive $CD8^+$ T cell pool can therefore be pathogenic or protective, depending on the nature of the infectious challenge.

In conclusion, our data provide direct evidence that a germline-encoded TCR segment can determine immune responsiveness to an exogenous peptide antigen, thereby extending the concept of *Ir* genes beyond the MHC. The unique features associated with this phenomenon may allow novel interventions to improve vaccine efficacy and limit immune pathology in humans, pending further studies to identify similar genetic associations between disease outcome and heritable components of the TCR.

AUTHOR CONTRIBUTIONS

NVBB, SG, KL, TMJ, LP, SLU, CF, DAP, JR and JTH designed the study. NVBB, SG, KL, TMJ, LP, SLU, KLM and CF conducted experiments. NVBB, SG, KL, TMJ, LP, SLU, KLM, CF, DAP, JR and JTH analyzed data and interpreted results. NVBB and SG drafted the manuscript. NVBB, SG, KL, TMJ, LP, SLU, CF, DAP, JR and JTH edited the manuscript.

ACKNOWLEDGEMENTS

We thank Marc Jenkins for helpful discussions, Vladimir Badovinac and Stanley Perlman for constructive comments on the manuscript, Lisa Hancox and Steven Moieffer for excellent technical support, and staff at the New York University Insectary Core and the University of Iowa Flow Cytometry Core. This work was supported by grants from the National Institutes of Health (NIH) to JTH (AI42767, AI85515, AI95178 and AI100527), the Australian Research Council (ARC) and the Australian National Health and Medical Research Council (NHMRC) to JR, and the Wellcome Trust to DAP (100326/Z/12/Z). NVBB was supported by NIH grant T32 AI007511, and SLU was supported by NIH grant T32 AI007343. SG is a Monash Senior Research Fellow, DAP is a Wellcome Trust Senior Investigator, and JR is an ARC Australian Laureate Fellow.

REFERENCES

Akue, A.D., Lee, J.Y., and Jameson, S.C. (2012). Derivation and maintenance of virtual memory CD8 T cells. J Immunol *188*, 2516-2523.

Amani, V., Vigario, A.M., Belnoue, E., Marussig, M., Fonseca, L., Mazier, D., and Renia, L. (2000). Involvement of IFN-gamma receptor-medicated signaling in pathology and anti-malarial immunity induced by Plasmodium berghei infection. Eur J Immunol *30*, 1646-1655.

Badovinac, V.P., and Harty, J.T. (2000). Adaptive immunity and enhanced CD8+ T cell response to Listeria monocytogenes in the absence of perforin and IFN-gamma. J Immunol *164*, 6444-6452.

Badovinac, V.P., Messingham, K.A., Jabbari, A., Haring, J.S., and Harty, J.T. (2005). Accelerated CD8+ T-cell memory and prime-boost response after dendritic-cell vaccination. Nat Med *11*, 748-756.

Behlke, M.A., Chou, H.S., Huppi, K., and Loh, D.Y. (1986). Murine T-cell receptor mutants with deletions of beta-chain variable region genes. PNAS *83*, 767-771.

Benacerraf, B. (1974). Editorial: Immune response genes. Scand J Immunol *3*, 381-386.

Benacerraf, B., and Germain, R.N. (1978). The immune response genes of the major histocompatibility complex. Immunol Rev *38*, 70-119.

Benacerraf, B., and McDevitt, H.O. (1972). Histocompatibility-linked immune response genes. Science *175*, 273-279.

Bendelac, A., Rivera, M.N., Park, S.H., and Roark, J.H. (1997). Mouse CD1-specific NK1 T cells: development, specificity, and function. Annu Rev Immunol *15*, 535-562.

Bettini, M.L., Bettini, M., Nakayama, M., Guy, C.S., and Vignali, D.A. (2013). Generation of T cell receptor-retrogenic mice: improved retroviral-mediated stem cell gene transfer. Nat Protoc *8*, 1837-1840.

Blattman, J.N., Antia, R., Sourdive, D.J., Wang, X., Kaech, S.M., Murali-Krishna, K., Altman, J.D., and Ahmed, R. (2002). Estimating the precursor frequency of naive antigen-specific CD8 T cells. J Exp Med *195*, 657-664.

Boubou, M.I., Collette, A., Voegtle, D., Mazier, D., Cazenave, P.A., and Pied, S. (1999). T cell response in malaria pathogenesis: selective increase in T cells carrying the TCR V(beta)8 during experimental cerebral malaria. Int Immunol *11*, 1553-1562.

Brewster, D.R., Kwiatkowski, D., and White, N.J. (1990). Neurological sequelae of cerebral malaria in children. Lancet *336*, 1039-1043.

Busch, D.H., and Pamer, E.G. (1999). T cell affinity maturation by selective expansion during infection. J Exp Med *189*, 701-710.

Busch, D.H., Pilip, I., and Pamer, E.G. (1998). Evolution of a complex T cell receptor repertoire during primary and recall bacterial infection. J Exp Med *188*, 61-70.

Cabaniols, J.P., Fazilleau, N., Casrouge, A., Kourilsky, P., and Kanellopoulos, J.M. (2001). Most alpha/beta T cell receptor diversity is due to terminal deoxynucleotidyl transferase. J Exp Med *194*, 1385-1390.

Chothia, C., Boswell, D.R., and Lesk, A.M. (1988). The outline structure of the T-cell alpha beta receptor. EMBO J *7*, 3745-3755.

Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr *50*, 760-763.

Dai, S., Huseby, E.S., Rubtsova, K., Scott-Browne, J., Crawford, F., Macdonald, W.A., Marrack, P., and Kappler, J.W. (2008). Crossreactive T cells spotlight the germline rules for alphabeta T cell-receptor interactions with MHC molecules. Immunity *28*, 324-334.

Davis, M.M., and Bjorkman, P.J. (1988). T-cell antigen receptor genes and T-cell recognition. Nature *334*, 395-402.

Davis, M.M., Krogsgaard, M., Huppa, J.B., Sumen, C., Purbhoo, M.A., Irvine, D.J., Wu, L.C., and Ehrlich, L. (2003). Dynamics of cell surface molecules during T cell recognition. Annu Rev Biochem *72*, 717-742.

Day, E.B., Guillonneau, C., Gras, S., La Gruta, N.L., Vignali, D.A., Doherty, P.C., Purcell, A.W., Rossjohn, J., and Turner, S.J. (2011). Structural basis for enabling T-cell receptor diversity within biased virus-specific CD8+ T-cell responses. PNAS *108*, 9536-9541.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System. http://www.pymol.org.

Doll, K.L., Pewe, L.L., Kurup, S.P., and Harty, J.T. (2016). Discriminating Protective from Nonprotective Plasmodium-Specific CD8+ T Cell Responses. J Immunol *196*, 4253-4262.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

Engwerda, C., Belnoue, E., Gruner, A.C., and Renia, L. (2005). Experimental models of cerebral malaria. Curr Top Microbiol Immunol *297*, 103-143.

Evans, P. (2006). Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr *62*, 72-82.

Fulton, R.B., Hamilton, S.E., Xing, Y., Best, J.A., Goldrath, A.W., Hogquist, K.A., and Jameson, S.C. (2015). The TCR's sensitivity to self peptide-MHC dictates the ability of naive CD8(+) T cells to respond to foreign antigens. Nat Immunol *16*, 107-117.

Garcia, K.C., Adams, J.J., Feng, D., and Ely, L.K. (2009). The molecular basis of TCR germline bias for MHC is surprisingly simple. Nat Immunol *10*, 143-147.

Godfrey, D.I., Uldrich, A.P., McCluskey, J., Rossjohn, J., and Moody, D.B. (2015). The burgeoning family of unconventional T cells. Nat Immunol *16*, 1114-1123.

Goldrath, A.W., and Bevan, M.J. (1999). Selecting and maintaining a diverse T-cell repertoire. Nature *402*, 255-262.

Gordon, E.B., Hart, G.T., Tran, T.M., Waisberg, M., Akkaya, M., Kim, A.S., Hamilton, S.E., Pena, M., Yazew, T., Qi, C.F.*, et al.* (2015). Targeting glutamine metabolism rescues mice from late-stage cerebral malaria. PNAS *112*, 13075-13080.

Gras, S., Burrows, S.R., Kjer-Nielsen, L., Clements, C.S., Liu, Y.C., Sullivan, L.C., Bell, M.J., Brooks, A.G., Purcell, A.W., McCluskey, J., and Rossjohn, J. (2009). The shaping of T cell receptor recognition by self-tolerance. Immunity *30*, 193-203.

Hafalla, J.C., Bauza, K., Friesen, J., Gonzalez-Aseguinolaza, G., Hill, A.V., and Matuschewski, K. (2013). Identification of targets of CD8(+) T cell responses to malaria liver stages by genome-wide epitope profiling. PLoS Pathog *9*, e1003303.

Haluszczak, C., Akue, A.D., Hamilton, S.E., Johnson, L.D., Pujanauski, L., Teodorovic, L., Jameson, S.C., and Kedl, R.M. (2009). The antigen-specific CD8+ T cell repertoire in unimmunized mice includes memory phenotype cells bearing markers of homeostatic expansion. J Exp Med *206*, 435-448.

Haque, A., Best, S.E., Unosson, K., Amante, F.H., de Labastida, F., Anstey, N.M., Karupiah, G., Smyth, M.J., Heath, W.R., and Engwerda, C.R. (2011). Granzyme B expression by CD8+ T cells is required for the development of experimental cerebral malaria. J Immunol *186*, 6148-6156.

Holst, J., Szymczak-Workman, A.L., Vignali, K.M., Burton, A.R., Workman, C.J., and Vignali, D.A. (2006). Generation of T-cell receptor retrogenic mice. Nat Protoc *1*, 406-417.

Horne-Debets, J.M., Karunarathne, D.S., Faleiro, R.J., Poh, C.M., Renia, L., and Wykes, M.N. (2016). Mice lacking Programmed cell death-1 show a role for CD8(+) T cells in long-term immunity against blood-stage malaria. Sci Rep *6*, 26210.

Howland, S.W., Poh, C.M., Gun, S.Y., Claser, C., Malleret, B., Shastri, N., Ginhoux, F., Grotenbreg, G.M., and Renia, L. (2013). Brain microvessel cross-presentation is a hallmark of experimental cerebral malaria. EMBO Mol Med *5*, 916-931.
Imai, T., Shen, J., Chou, B., Duan, X., Tu, L., Tetsutani, K., Moriya, C., Ishida, H., Hamano, S., Shimokawa, C.*, et al.* (2010). Involvement of CD8+ T cells in protective immunity against murine blood-stage infection with Plasmodium yoelii 17XL strain. Eur J Immunol *40*, 1053-1061.

Jenkins, M.K., and Moon, J.J. (2012). The role of naive T cell precursor frequency and recruitment in dictating immune response magnitude. J Immunol *188*, 4135-4140.

Jerne, N.K. (1971). The somatic generation of immune recognition. Eur J Immunol *1*, 1-9.

Kabsch, W. (2010). XDS. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Kedl, R.M., Rees, W.A., Hildeman, D.A., Schaefer, B., Mitchell, T., Kappler, J., and Marrack, P. (2000). T cells compete for access to antigen-bearing antigen-presenting cells. J Exp Med *192*, 1105-1113.

Kjer-Nielsen, L., Clements, C.S., Brooks, A.G., Purcell, A.W., McCluskey, J., and Rossjohn, J. (2002). The 1.5 A crystal structure of a highly selected antiviral T cell receptor provides evidence for a structural basis of immunodominance. Structure *10*, 1521-1532.

Lefranc, M.P. (2003). IMGT databases, web resources and tools for immunoglobulin and T cell receptor sequence analysis, http://imgt.cines.fr. Leukemia *17*, 260-266.

Mach, N., Gillessen, S., Wilson, S.B., Sheehan, C., Mihm, M., and Dranoff, G. (2000). Differences in dendritic cells stimulated in vivo by tumors engineered to secrete granulocyte-macrophage colony-stimulating factor or Flt3-ligand. Cancer Res *60*, 3239-3246.

Malherbe, L., Hausl, C., Teyton, L., and McHeyzer-Williams, M.G. (2004). Clonal selection of helper T cells is determined by an affinity threshold with no further skewing of TCR binding properties. Immunity *21*, 669-679.

Mandl, J.N., Monteiro, J.P., Vrisekoop, N., and Germain, R.N. (2013). T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. Immunity *38*, 263-274.

Mariotti-Ferrandiz, E., Pham, H.P., Dulauroy, S., Gorgette, O., Klatzmann, D., Cazenave, P.A., Pied, S., and Six, A. (2016). A TCRbeta Repertoire Signature Can Predict Experimental Cerebral Malaria. PLoS One *11*, e0147871.

Marshak, A., Doherty, P.C., and Wilson, D.B. (1977). The control of specificity of cytotoxic T lymphocytes by the major histocompatibility complex (AG-B) in rats and identification of a new alloantigen system showing no AG-B restriction. J Exp Med *146*, 1773-1790.

McDevitt, H.O., and Chinitz, A. (1969). Genetic control of the antibody response: relationship between immune response and histocompatibility (H-2) type. Science *163*, 1207-1208.

Miles, J.J., Douek, D.C., and Price, D.A. (2011). Bias in the alphabeta T-cell repertoire: implications for disease pathogenesis and vaccination. Immunol Cell Biol *89*, 375-387.

Moon, J.J., Chu, H.H., Pepper, M., McSorley, S.J., Jameson, S.C., Kedl, R.M., and Jenkins, M.K. (2007). Naive CD4(+) T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. Immunity *27*, 203-213.

Neller, M.A., Ladell, K., McLaren, J.E., Matthews, K.K., Gostick, E., Pentier, J.M., Dolton, G., Schauenburg, A.J., Koning, D., Fontaine Costa, A.I.*, et al.* (2015). Naive CD8(+) T-cell precursors display structured TCR repertoires and composite antigen-driven selection dynamics. Immunol Cell Biol *93*, 625-633.

Obar, J.J., Khanna, K.M., and Lefrancois, L. (2008). Endogenous naive CD8+ T cell precursor frequency regulates primary and memory responses to infection. Immunity *28*, 859-869.

Owens, T., and Zeine, R. (1989). The cell biology of T-dependent B cell activation. Biochem Cell Biol *67*, 481-489.

Pham, N.L., Badovinac, V.P., and Harty, J.T. (2009). A default pathway of memory CD8 T cell differentiation after dendritic cell immunization is deflected by encounter with inflammatory cytokines during antigen-driven proliferation. J Immunol *183*, 2337-2348.

Price, D.A., Brenchley, J.M., Ruff, L.E., Betts, M.R., Hill, B.J., Roederer, M., Koup, R.A., Migueles, S.A., Gostick, E., Wooldridge, L.*, et al.* (2005). Avidity for antigen shapes clonal dominance in CD8+ T cell populations specific for persistent DNA viruses. J Exp Med *202*, 1349-1361.

Quigley, M.F., Almeida, J.R., Price, D.A., and Douek, D.C. (2011). Unbiased molecular analysis of T cell receptor expression using template-switch anchored RT-PCR. Curr Protoc Immunol *Chapter 10*, Unit10 33.

Quigley, M.F., Greenaway, H.Y., Venturi, V., Lindsay, R., Quinn, K.M., Seder, R.A., Douek, D.C., Davenport, M.P., and Price, D.A. (2010). Convergent recombination shapes the clonotypic landscape of the naive T-cell repertoire. PNAS *107*, 19414-19419.

Rai, D., Pham, N.L., Harty, J.T., and Badovinac, V.P. (2009). Tracking the total CD8 T cell response to infection reveals substantial discordance in magnitude and kinetics between inbred and outbred hosts. J Immunol *183*, 7672-7681.

Read, R.J. (2001). Pushing the boundaries of molecular replacement with maximum likelihood. Acta Crystallogr D Biol Crystallogr *57*, 1373-1382.

Rossjohn, J., Gras, S., Miles, J.J., Turner, S.J., Godfrey, D.I., and McCluskey, J. (2015). T cell antigen receptor recognition of antigen-presenting molecules. Annu Rev Immunol *33*, 169-200.

Savage, P.A., Boniface, J.J., and Davis, M.M. (1999). A kinetic basis for T cell receptor repertoire selection during an immune response. Immunity *10*, 485-492.

Scott-Browne, J.P., White, J., Kappler, J.W., Gapin, L., and Marrack, P. (2009). Germline-encoded amino acids in the alphabeta T-cell receptor control thymic selection. Nature *458*, 1043-1046.

Struik, S.S., and Riley, E.M. (2004). Does malaria suffer from lack of memory? Immunol Rev *201*, 268-290.

Townsend, A.R., Gotch, F.M., and Davey, J. (1985). Cytotoxic T cells recognize fragments of the influenza nucleoprotein. Cell *42*, 457-467.

Turner, S.J., Doherty, P.C., McCluskey, J., and Rossjohn, J. (2006). Structural determinants of T-cell receptor bias in immunity. Nat Rev Immunol *6*, 883-894.

Valkenburg, S.A., Quinones-Parra, S., Gras, S., Komadina, N., McVernon, J., Wang, Z., Halim, H., Iannello, P., Cole, C., Laurie, K.*, et al.* (2013). Acute emergence and reversion of influenza A virus quasispecies within CD8+ T cell antigenic peptides. Nat Commun *4*, 2663.

van der Heyde, H.C., Manning, D.D., Roopenian, D.C., and Weidanz, W.P. (1993). Resolution of blood-stage malarial infections in CD8+ cell-deficient beta 2-m0/0 mice. J Immunol *151*, 3187-3191.

van der Merwe, P.A., and Dushek, O. (2011). Mechanisms for T cell receptor triggering. Nat Rev Immunol *11*, 47-55.

Van Rhijn, I., Godfrey, D.I., Rossjohn, J., and Moody, D.B. (2015). Lipid and small-molecule display by CD1 and MR1. Nat Rev Immunol *15*, 643-654.

Vinetz, J.M., Kumar, S., Good, M.F., Fowlkes, B.J., Berzofsky, J.A., and Miller, L.H. (1990). Adoptive transfer of CD8+ T cells from immune animals does not transfer immunity to blood stage Plasmodium yoelii malaria. J Immunol *144*, 1069-1074.

Wilson, A., Marechal, C., and MacDonald, H.R. (2001). Biased V beta usage in immature thymocytes is independent of DJ beta proximity and pT alpha pairing. J Immunol *166*, 51-57.

Yanez, D.M., Manning, D.D., Cooley, A.J., Weidanz, W.P., and van der Heyde, H.C. (1996). Participation of lymphocyte subpopulations in the pathogenesis of experimental murine cerebral malaria. J Immunol *157*, 1620-1624.

Young, A.C., Zhang, W., Sacchettini, J.C., and Nathenson, S.G. (1994). The three-dimensional structure of H-2Db at 2.4 A resolution: implications for antigen-determinant selection. Cell *76*, 39-50.

Zhao, J., Zhao, J., and Perlman, S. (2009). De novo recruitment of antigen-experienced and naive T cells contributes to the long-term maintenance of antiviral T cell populations in the persistently infected central nervous system. J Immunol *183*, 5163-5170.

Zinkernagel, R.M. (1978). Thymus and lymphohemopoietic cells: their role in T cell maturation in selection of T cells' H-2-restriction-specificity and in H-2 linked Ir gene control. Immunol Rev *42*, 224-270.

FIGURE LEGENDS

Figure 1: $GAP50_{40-48}$-specific $CD8^+$ T cells express biased TCRs

(A, B) Mice were primed with peptide-coated DCs and boosted on day 7 with recombinant *L. monocytogenes* expressing the same epitope. (A) The TCR Vβ repertoire was assessed using splenocytes isolated from immunized mice by staining with the indicated tetramers and Vβ-specific antibodies. Results are shown for two mice for each antigen specificity. Data represent four independent experiments (n = 5 mice/group). Cumulative data are shown in Supplementary Figure 1B. (B) Extreme focusing of the $GAP50_{40-48}$-specific repertoire does not reflect the overall frequency of $V\beta8.1/8.2^+$ clonotypes in the $CD8^+$ T cell compartment. Data represent three independent experiments (n = 4–5 mice/group). (C, D) Confirmation of TCR bias at the transcriptional level. Sequences were derived from $GAP50_{40-48}$-specific $CD8^+$ T cells isolated from the spleens of mice previously immunized with DC-LM-$GAP50_{40-48}$ (total = 316 molecular clones). (C) Population-level analysis revealed an extreme bias towards Vβ8.1. (D) TCR Vα sequence distributions from three different mice are depicted. No obvious bias was apparent in the Vα repertoire. (E) Origin of CDR3β amino acids expressed relative to the total number of molecular $V\beta8.1^+$ clones (n = 5 mice). Vβ: red; Dβ: green; Jβ: purple; N segments: white. (See also Figures S1–S3 and Table S1).

Figure 2: TCR interactions focus on the C-terminal region of the $GAP50_{40-48}$ peptide

(A) The structure of the $GAP50_{40-48}$ peptide in complex with $H\text{-}2D^b$. The $GAP50_{40-48}$ peptide is represented as black sticks in the antigen-binding cleft of $H\text{-}2D^b$ (white cartoon). The P2-Q and P9-L anchor residues are buried within the antigen-binding cleft, and P5-N acts as a secondary anchor residue forming hydrogen bonds (black dashes) with Gln97 (white stick) at the base of the antigen-binding cleft of $H\text{-}2D^b$ (white cartoon). (B) A panel of peptide mutants was generated by replacing amino acids with Ala or Ser at selected positions (red). Anchor residues are underlined. (C) Splenocytes isolated from $GAP50_{40-48}$-immunized mice were restimulated for 5 h *in vitro* using *wt* or mutant peptides. The production of IFNγ induced by stimulation with the different mutant

peptides is expressed relative to the production of IFNγ induced by stimulation with the *wt* peptide. Data represent two independent experiments performed in triplicate. Bars represent mean ± SD. Significance was assessed using a one-way ANOVA (****$p < 0.0001$). (D) The effect of peptide substitutions on recognition of the H-2D^b-GAP50$_{40-48}$ complex by specific CD8$^+$ T cell clones. The surface of H-2D^b is shown in white. Peptide residues that were not mutated or did not affect IFNγ production are shown in gray. Green represents up to a 25% decrease in T cell activation; orange represents up to a 70% decrease in T cell activation; red represents a total abrogation of T cell activation. (See also Figure S4 and Table S2).

Figure 3: GAP50$_{40-48}$-specific TCR bias is peptide-driven

(A) The crystal structure of the NB1 TCR (α chain: pink cartoon; β chain: purple cartoon) bound to the GAP50$_{40-48}$ peptide (black sticks) presented by H-2D^b (heavy chain: white cartoon; β2m, wheat cartoon). (B) The CDR loop contribution to the BSA is represented as teal (CDR1α), green (CDR2α), purple (CDR3α), red (CDR1β), orange (CDR2β) or yellow (CDR3β). Black spheres represent the center of mass for the Vα and Vβ. (C–E) The contribution (%) of the CDR loops to the interaction with H-2D^b-GAP50$_{40-48}$ (C), GAP50$_{40-48}$ (D) and H-2D^b (E). (F–K) Contact residues between H-2D^b-GAP50$_{40-48}$ and CDR1α (F), CDR2α (G), CDR3α (H), CDR1β (I, J), CDR2β (J) and CDR3β (K). The CDR color scheme is maintained though panels F–K. Residues that make contacts with the GAP50$_{40-48}$ peptide or H-2D^b (white) are depicted as black and white sticks, respectively. The H-2D^b molecule is represented as white cartoon, and hydrogen bonds are shown as black dashes. (See also Figure S5 and Tables S2–S5).

Figure 4: Binding affinities of NB1 TCR mutants for H-2D^b-GAP50$_{40-48}$

(A) Surface plasmon resonance (SPR) analysis of NB1 *wt* (black) and the indicated mutant (colored) TCRs across a range of H-2D^b-GAP50$_{40-48}$ concentrations up to a maximum of 200 μM. Representative SPR binding curves are shown for the NB1 *wt* TCR (black) and NB1 mutant TCRs with the following β chain substitutions: Y31A (green), D30A (red), Y57A (pink), V58A (orange) and Y31A (blue). Each NB1 mutant TCR was tested in parallel with the NB1 *wt* TCR. (B) Summary table representing

equilibrium binding affinities of NB1 TCR mutants for H-2D^{b}-GAP50$_{40\text{-}48}$. NB: no binding; RU: response units. Data represent two independent experiments performed in duplicate (mean ± SEM). (See also Tables S6A and S6B).

Figure 5: The naive GAP50$_{40\text{-}48}$-specific CD8^{+} T cell repertoire is extremely large

(A) CD8^{+} T cells specific for OVA$_{257\text{-}264}$, GP$_{33\text{-}41}$, GAP50$_{40\text{-}48}$, S20$_{318\text{-}326}$ and TRAP$_{130\text{-}138}$ were enriched from the spleens and macroscopic lymph nodes of naive C57Bl/6 mice and used to calculate the final numbers of tetramer^{+} cells indicated in each plot. Target cells were identified as CD8^{+} CD90.2^{+} CD11b^{-} CD11c^{-} B220^{-} tetramer(APC)$^{+}$ tetramer(PE)$^{+}$. Depicted numbers of naive precursors were calculated as mean ± SD from 11 mice for the GAP50$_{40\text{-}48}$-specific repertoire. Right panel: Vβ8.1/8.2 expression on naive GAP50$_{40\text{-}48}$-specific CD8^{+} T cells. (B) Sizes of the naive repertoires specific for various CD8^{+} T cell epitopes. The number of naive GAP50$_{40\text{-}48}$-specific precursors is represented as a range from 2,700–3,250 cells (total n = 11 mice). (C, D) Phenotypic characterization of naive GAP50$_{40\text{-}48}$-specific CD8^{+} T cells. (C) Expression of CD44 on naive GAP50$_{40\text{-}48}$-specific CD8^{+} T cells was compared with CD8^{+} T cells obtained from naive Tg OT-I and P14 mice. (D) Expression of CD5 on naive GAP50$_{40\text{-}48}$-specific CD8^{+} T cells was compared with total naive, non-GAP50$_{40\text{-}48}$-specific CD8^{+} T cells (left panel) and with naive Tg OT-I cells (right panel). Data represent two independent experiments (n = 4 mice/group). Bars depict mean ± SD.

Figure 6: Vβ8.1 is required for the generation of GAP50$_{40\text{-}48}$-specific CD8^{+} T cells and the development of ECM

(A, B) C57Bl/6 and C57/L mice were infected with 1,000 *P. berghei* ANKA sporozoites. H-2D^{b}-GAP50$_{40\text{-}48}$ tetramer^{+} cells were quantified in the blood (A) and brains (B) at day 7 post-infection (representative plots: left; summary graph: right). (C) Magnitude of the activated CD8^{+} T cell response at day 7 post-infection, expressed as % CD11a^{hi} CD8^{lo} of total CD8^{+} T cells (representative plots: left; summary graph: right). Naive control values were subtracted from individual values. Data represent two independent experiments (n = 3 mice/group). Bars depict mean ± SD. Significance was assessed using an unpaired, two-tailed t test (**p < 0.01, ***p < 0.001). (D) Survival curves for

C57Bl/6 mice and C57/L mice after infection with *P. berghei* ANKA. Data represent three independent experiments (n = 5 mice/group). Significance was assessed using the Mantel-Cox log rank test (***p = 0.001). (E) C57Bl/6 and CB6F1 mice were infected with 10^6 *P. berghei*-parasitized red blood cells (pRBCs). Mice were monitored for the development of ECM symptoms and scored for survival over a period of two weeks. Data represent three independent experiments (n = 5 mice/group). Significance was assessed using the Mantel-Cox log rank test (**p = 0.0025). (F) $GAP50_{40-48}$-specific $CD8^+$ T cells were tetramer-enriched from naive C57Bl/6 and CB6F1 mice and quantified to estimate repertoire size. Cumulative results are shown from two independent experiments (total n = 4 mice/group). Bars depict mean ± SEM. Significance was assessed using an unpaired, two-tailed t test (****p <0.0001). (G) C57Bl/6 and CB6F1 mice were infected with 10^6 pRBCs. The $GAP50_{40-48}$-specific $CD8^+$ T cell response was followed in the blood of infected mice using H-$2D^b$-$GAP50_{40-48}$ tetramers. All C57Bl/6 mice succumbed to ECM by days 6 and 7, while all CB6F1 mice survived. Data represent two independent experiments (n = 5 mice/group). Bars depict mean ± SD. Significance was assessed using an unpaired, two-tailed t test (****p < 0.0001). (H, I) C57Bl/6 and CB6F1 mice were injected IV with 5 x10^5 $GAP50_{40-48}$ peptide-coated DCs and then infected 7 days later with recombinant attenuated LM-$GAP50_{40-48}$. At a memory time point after LM infection, DC-LM-$GAP50_{40-48}$-immunized or non-immunized mice were infected with 10^6 pRBCs. Survival curves are shown for C57Bl/6 mice (H) and CB6F1 mice (I). Data represent two independent experiments (n = 5 mice/group). Significance was assessed using the Mantel-Cox log rank test (p = 0.0993 in panel D; **p = 0.0018 in panel E).

Figure 7: Large numbers of $GAP50_{40-48}$-specific $CD8^+$ T cells control primary *L. monocytogenes* infection

(A) BALB/c mice were infected IV with 5x10^3 CFU of *wt* LM or recombinant LM-$GAP50_{40-48}$. Bacterial burden was measured in the spleen on day 5. Data represent two independent experiments (n = 4 mice/group). Bars depict mean ± SD. (B) C57Bl/6 mice were infected IV with 10^4 CFU of *wt* LM or recombinant LM-$GAP50_{40-48}$. Bacterial burden was measured in the spleen on day 5. Cumulative results are shown from three

independent experiments (total n = 14–15 mice/group). Bars depict mean ± SEM. Significance was assessed using an unpaired, two-tailed t test (****p<0.0001). (C) Representative flow cytometry plots showing H-2D^{b}-GAP50$_{40\text{-}48}$ tetramer staining of splenocytes isolated from infected C57Bl/6 mice at day 5 post-infection.

METHODS

Contact for Reagent and Resource Sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, John T. Harty (john-harty@uiowa.edu).

Experimental Model and Subject Details

C57Bl/6 (H-2^b), BALB/c (H-2^d) and CB6F1 (H-2^{bxd}) mice were purchased from the National Cancer Institute (Frederick, MD). BALB.b (H-2^b) and C57/L (H-2^b) mice were purchased from the Jackson Laboratory (Bar Harbor, ME). All mice were housed with appropriate biosafety containment at the University of Iowa Animal Care Unit. The animals were treated and handled in accordance with guidelines established by the Institutional Animal Care and Use Committee. All experiments were performed using female mice, aged 6–8 weeks. *P. berghei* ANKA clone 234-parasitized *Anopheles stephensi* mosquitos were produced in house or obtained from New York University. For ECM infection, mice were injected intravenously (IV) with 10^3 viable sporozoites or intraperitoneally (IP) with 10^6 pRBCs. pRBCs were obtained from mice infected with a frozen stock 4–5 days previously. Numbers of transferred pRBCs were estimated by counting infected cells in Giemsa-stained blood smears. For virulent LM infection (*wt* or LM-GAP$50_{40\text{-}48}$), BALB/c mice were infected IV with $5x10^3$ CFU, and C57Bl/6 mice were infected IV with 10^4 CFU. Bacterial burden in the spleen was measured according to a previously published protocol (Badovinac and Harty, 2000).

Method Details

DC-LM immunizations

For the purpose of TCR clonotyping, mice were immunized using a DC-LM prime-boost regimen (Badovinac et al., 2005). FMS-like tyrosine kinase-3 ligand (Flt-3L)-induced splenic DCs were LPS-matured, isolated after collagenase/DNAse digestion (Pham et al., 2009), and incubated with the following peptides for 2 h at a final concentration of 2 μM: $TRAP_{130\text{-}138}$ (SALLNUDNL), $S20_{318\text{-}326}$ (VNYSFLYLF) or $GAP50_{40\text{-}48}$ (SQLLNAKYL). $CD11c^+$ cells were then enriched using anti-CD11c MicroBeads (Miltenyi Biotec). A total

of $5x10^5$ purified DCs in saline were used to prime each mouse IV. After 7 days, mice were boosted IV with 10^7 CFU of recombinant attenuated LM-TRAP$_{130-138}$, LM-S20$_{318-326}$ or LM-GAP50$_{40-48}$. All recombinant attenuated LM strains were both *actA*- and *inlB*-deficient. The number of injected bacteria was verified by plating on tryptic soy agar supplemented with 50 µg/ml streptomycin.

TCR clonotyping

C57Bl/6 mice were immunized against GAP50$_{40-48}$ using the DC-LM prime-boost regimen. After 5 weeks, splenocytes were isolated and stained with LIVE/DEAD Fixable Aqua and anti-CD16/CD32 (FCR4G8; Thermo Fisher Scientific). Antigen-specific cells were labeled using a BV421-conjugated version of the GAP50$_{40-48}$ tetramer. Lineage markers were identified using anti-CD3-Cy5PE (17A2; BioLegend), anti-CD4-Alexa Fluor 700 (RM4-5; BD Biosciences) and anti-CD8-BV711 (53-6.7; BioLegend). Viable $CD3^+$ $CD4^-$ $CD8^+$ GAP50$_{40-48}$ $tetramer^+$ cells were then sorted at >98% purity using a custom-built 20-parameter FACSAria II (BD Biosciences). TCR clonotyping was performed from 5,000 sorted cells per mouse using a template-switch anchored RT-PCR (Quigley et al., 2011). Amplicons were subcloned, sampled, sequenced and analyzed as described previously (Price et al., 2005). In all cases, TCR nomenclature was translated from the IMGT database via web-based alignment of molecular transcripts (http://www.imgt.org).

Tetramer-based $CD8^+$ T cell enrichment

The size of the naive antigen-specific $CD8^+$ T cell repertoire was quantified using a modified version of a previously published protocol (Moon et al., 2007). Spleen, popliteal, inguinal, brachial, axillary and cervical lymph nodes were isolated from naive mice and processed to a single cell suspension. Prepared cells were then labeled with an equimolar mixture of PE- and APC-conjugated tetramers representing each defined epitope. After incubation for 1 h at 4°C, cells were surface stained for CD8, CD90.2 and exclusion markers (B220, CD11b and CD11c). Labeled cells were subsequently captured and enumerated using anti-PE and anti-APC magnetic beads with an AutoMACS separator (Miltenyi Biotec).

Peptide mutants

Mutant $GAP50_{40-48}$ peptides were generated by replacing individual amino acids with Ala or Ser (for P6-Ala), barring the anchor residues at P5 and P9 (BioSynthesis Inc.). Splenocytes were isolated from DC-LM-$GAP50_{40-48}$-immunized mice during the memory phase (>40 days post-immunization) and restimulated for 5 h in the presence of brefeldin A with either *wt* or mutant peptides at a final concentration of 500 nM. All restimulations were performed in triplicate. IFNγ-producing $CD8^+$ T cells were identified using a standard intracellular cytokine staining protocol (Doll et al., 2016). Data are expressed relative to the percentage of $CD8^+$ T cells producing IFNγ in response to the *wt* peptide.

Generation of TCRβ retrogenic mice

TCRβ retrogenic mice were generated on a C57Bl/6 background using retrovirus-mediated stem cell gene transfer (Bettini et al., 2013; Holst et al., 2006). The relevant antigen-specific TCRβ sequences were obtained from $CD8^+$ T cell clones derived from mice immunized with $GAP50_{40-48}$ ($V\beta8.1^+$) or $TRAP_{130-138}$ ($V\beta9^+$).

Isolation of mononuclear cells from the brains of P. berghei ANKA-infected mice

Mononuclear cells were isolated from brain tissue using a previously published protocol with minor modifications (Zhao et al., 2009). Brains harvested after intravascular exclusion were processed and digested with collagenase D (1 mg/ml; Roche) and DNase (0.1 mg/ml; Sigma-Aldrich) for 30 min at 37°C. Dissociated brain tissue was then passed through a 70 µm nylon mesh cell strainer, spun down, resuspended and centrifuged for 20 min over 37% Percoll. Mononuclear cells were collected from the pellet, and RBCs were lysed using 1X VitaLyse (CMDG).

Protein expression, purification and crystallization

The α and β chains of the TCR clones NB1, KL1 and KL4 were expressed separately as inclusion bodies and refolded with engineered disulfide linkages in the constant domains as described previously (Day et al., 2011). Ala substitutions in the $H\text{-}2D^b$ and

NB1 TCR sequences were introduced using site-directed mutagenesis. Soluble *wt* and mutant H-2D^b heterodimers containing the GAP50$_{40-48}$ peptide were prepared as described previously (Day et al., 2011). Crystals of H-2D^b-GAP50$_{40-48}$ (5 mg/ml) or the NB1 TCR in complex with H-2D^b-GAP50$_{40-48}$ (6 mg/ml) in 10 mM Tris-HCl (pH 8) and 150 mM NaCl were grown by the hanging-drop, vapor-diffusion method at 20°C with a protein/reservoir drop ratio of 1:1. Crystals of free H-2D^b-GAP50$_{40-48}$ were obtained in 0.1 M Tris-HCl (pH 8.5), 0.2 M lithium sulfate and 25–30% PEG8000. Crystals of the NB1 TCR in complex with H-2D^b-GAP50$_{40-48}$ were obtained in 18% PEG3350, 2% ethylene glycol, 0.2 M $CaCl_2$ and 0.1 M HEPES (pH 7.2).

Data collection and structure determination

Crystals were soaked in a cryoprotectant solution containing mother liquor with the PEG8000 concentration increased to 30% (v/v) and then flash frozen in liquid nitrogen. Data were collected on the MX2 beamline at the Australian Synchrotron using the ADSC-Quantum 315r CCD detectors (at 100K). Data were processed with XDS software (Kabsch, 2010) and scaled using SCALA software (Evans, 2006) from the CCP4 suite (Collaborative Computational Project, 1994). The NB1 TCR structure was determined by molecular replacement using the Phaser program (Read, 2001) with the LC13 TCR as the search model (Protein Data Bank accession number, 1KGC) (Kjer-Nielsen et al., 2002), and the H-2D^b structure was determined by molecular replacement using the Phaser program (Read, 2001) with H-2D^b as the search model (Protein Data Bank accession number, 4L8D) (Valkenburg et al., 2013). Manual model building was conducted using Coot software (Emsley and Cowtan, 2004), followed by maximum-likelihood refinement with Buster. The NB1 TCR was numbered according to the IMGT unique numbering system (Lefranc, 2003). The final models were validated using the Protein Data Bank validation web site, and the final refinement statistics are summarized in Table S2. All molecular graphics representations were created using PyMol (DeLano, 2002).

Surface plasmon resonance

Surface plasmon resonance experiments were conducted at 25°C on a BIAcore 3000 instrument using 10 mM Tris-HCl (pH 8) and 150 mM NaCl supplemented with 1% BSA and 0.005% surfactant P20. The TCRs (*wt* or mutant) were immobilized on research-grade CM5 chips via standard amine coupling. All experiments were carried out at least twice in duplicate as described previously (Gras et al., 2009) across a H-2D^{b}-GAP50$_{40\text{-}48}$ (*wt* or mutant) concentration range of 0.78–200 μM. Data were analyzed using BIAevaluation version 3.1 with the 1:1 Langmuir binding model.

Quantification and Statistical Analysis

Statistical differences between two study groups were evaluated using an unpaired, two-tailed t test. Statistical differences between more than two study groups were evaluated using a one-way ANOVA with Tukey's multiple comparison post-hoc test. Bar graphs display mean ± SD for representative experiments and mean ± SEM for combined experiments. Statistical significance was assigned as *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ and ****$p < 0.0001$. Statistical analyses were performed using Prism 7 software (GraphPad).

Data and Software Availability

Structural coordinates were submitted to the Protein Data Bank with accession codes 5WLI for H-2D^{b}-GAP50$_{40\text{-}48}$ and 5WLG for the NB1 TCR in complex with H-2D^{b}-GAP50$_{40\text{-}48}$.

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-mouse CD8α (clone 53-6.7)	BioLegend	Cat# 100725
Anti-mouse CD11a (clone M17/4)	BioLegend	Cat# 101107
Anti-mouse CD90.2 (clone 30-H12)	BioLegend	Cat# 105306
Mouse Vβ screening Panel (Vβ 2, 3, 4, 5.1, 5.2, 6, 7, 8.1/8.2, 8.3, 9, 10[b], 11, 12, 13, 14 and 17[a])	BD Pharmingen	Cat# 557004
Anti-mouse IFNγ (clone XMG1.2)	BioLegend	Cat# 505809
Anti-mouse CD5 (clone 53-7.3)	BioLegend	Cat# 100629
Anti-mouse CD44 (clone IM7)	eBioscience	Cat# 11-0441-82
Anti-mouse CD11b (clone M1/70)	eBioscience	Car# 45-0112-82
Anti-mouse CD11c (clone N418)	eBioscience	Cat# 45-0114-82
Anti-mouse B220 (clone RA3-6B2)	eBioscience	Cat# 45-0452-82
Anti-mouse CD3 (clone 17A2)	BioLegend	Cat# 100309
Anti-mouse CD4 (clone RM4-5)	BD Bioscience	Cat# 100536
Anti-CD16/32 (clone FCR4G8)	Thermo Fisher Scientific	Cat# MFCR00
Tetramers		
H-2D^{b}-TRAP$_{130-138}$	*Prepared In house*	NA
H-2K^{b}-S20$_{318-326}$	*Prepared In house*	NA
H-2D^{b}-GAP50$_{40-48}$	*Prepared In house*	NA
H-2K^{b}-OVA$_{257-264}$	*Prepared In house*	NA
H-2D^{b}-GP$_{33-41}$	*Prepared In house*	NA
Other reagents used for flow cytometry		
LIVE/DEAD® Fixable Aqua	Thermo Fisher Scientific	Cat# L34957
Brilliant Violet BV421™ Streptavidin	BioLegend	Cat# 405225

Chemical and Peptides		
RNAlater® Stabilization Solution	Thermo Fisher Scientific	Cat# AM7020
RNaseOUT™ Recombinant Ribonuclease Inhibitor	Thermo Fisher Scientific	Cat# 10777019
10mM dNTP mix	Thermo Fisher Scientific	Cat# 18427013
SYBR® Gold Nucleic Acid Gel Stain	Thermo Fisher Scientific	Cat# S-11494
S.O.C. medium	Thermo Fisher Scientific	Cat# 15544034
SYBR® Safe DNA Gel Stain	Thermo Fisher Scientific	Cat# S33102
TrackIt™ Cyan/Orange Loading Buffer	Thermo Fisher Scientific	Cat# 10482-028
GeneRuler 1kb DNA ladder, ready-to-use	Thermo Fisher Scientific	Cat# SM0313
LPS	Sigma	Cat# L-8274
Brefeldin A (1,000x)	BioLegend	Cat# 420601
Collagenase Type II	Gibco	Cat# 17101-015
DNase	Sigma-Aldrich	Cat# D4513-1VL
Percoll	GE Healthcare	Cat# 17-0891-01
VitaLyse	CMDG	Cat# WBL0100
$TRAP_{130-138}$ peptide	BioSynthesis Inc	Lot# T7825-2
$S20_{318-326}$ peptide	BioSynthesis Inc	Lot# T7825-3
$GAP50_{40-48}$ peptide SQLLNAKYL	BioSynthesis Inc	Lot# T7982
$GAP50_{40-48}$ peptide mutant AQLLNAKYL	BioSynthesis Inc	Lot# P1168-1
$GAP50_{40-48}$ peptide mutant SALLNAKYL	BioSynthesis Inc	Lot# P1168-2
$GAP50_{40-48}$ peptide mutant SQALNAKYL	BioSynthesis Inc	Lot# P1168-3
$GAP50_{40-48}$ peptide mutant SQLANAKYL	BioSynthesis Inc	Lot# P1168-4

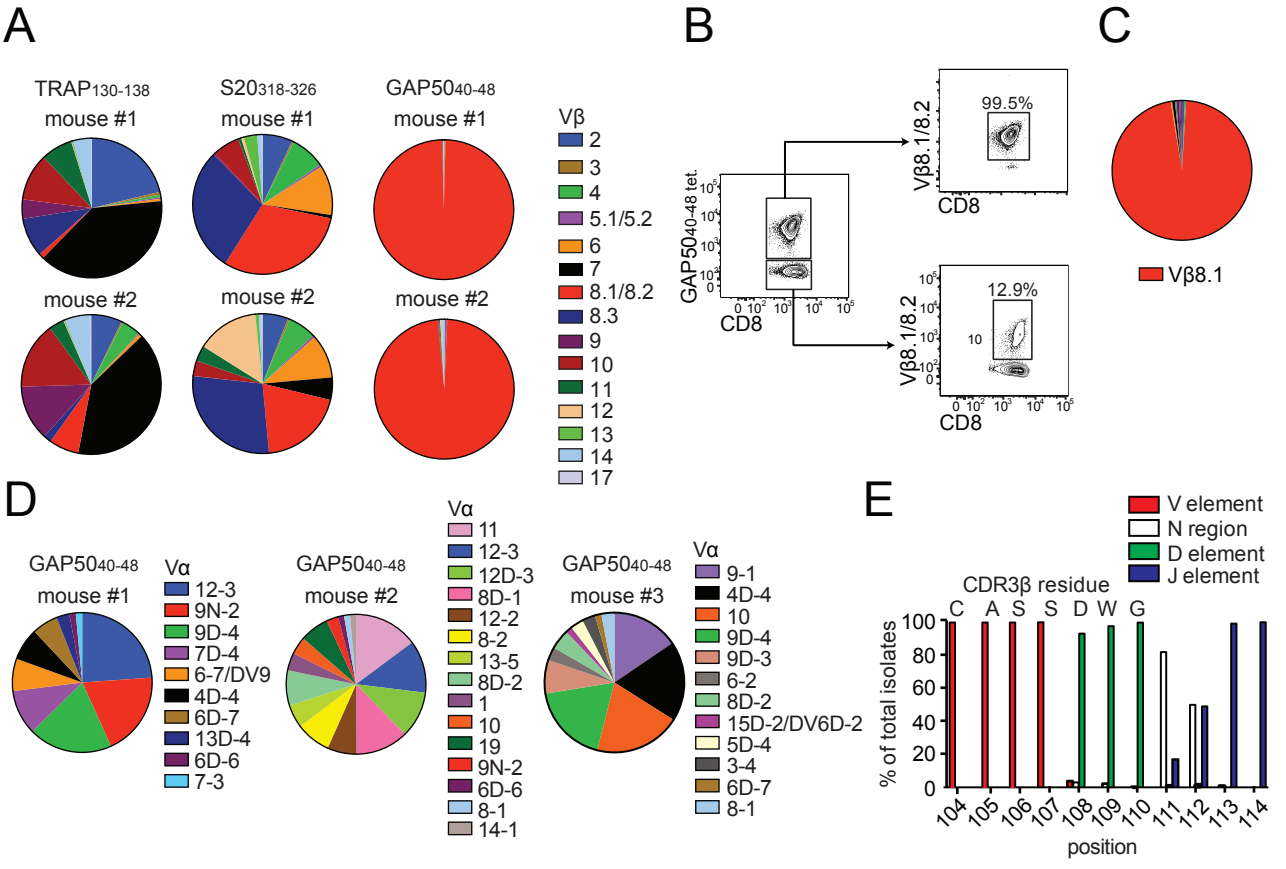
$GAP50_{40-48}$ peptide mutant SQLLNAAYL	BioSynthesis Inc	Lot# P1168-5
$GAP50_{40-48}$ peptide mutant SQLLNAKAL	BioSynthesis Inc	Lot# P1168-6
NB1, KL1 and KL4 TCR constructs	DNA from Genscript	NA
$H\text{-}2D^b$-$GAP50_{40-48}$, NB1, KL1 and KL4 TCR proteins	In house	Ref: (Day et al., 2011)
Commercial Assays		
anti-CD11c MicroBeads	Miltenyi Biotec	Cat# 130-108-338
Anti-PE MicroBeads	Miltenyi Biotec	Cat# 130-048-801
Anti-APC MicroBeads	Miltenyi Biotec	Cat# 130-090-855
µMACS mRNA isolation kit	Milteniyi Biotec	Cat# 130-075-201
SMARTer® RACE 5'/3' Kit	Clontech	Cat# 634859
SuperScript® II Reverse Transcriptase	Thermo Fisher Scientific	Cat# 18064014
TOPO® TA Cloning® Kit for Sequencing with One Shot® MAX Efficiency® DH5α-T1R E. coli	Thermo Fisher Scientific	Cat# K459540
Deposited Data		
$H\text{-}2D^b$-$GAP50_{40-48}$ structure	PDB web site	PDB: 5WLI
NB1 TCR-$H\text{-}2D^b$-$GAP50_{40-48}$ structure	PDB web site	PDB: 5WLG
Experimental models: Cell lines		
Mouse: primary lymphocytes	This manuscript	NA
Flt-3L tumor cell line	Ref: (Mach et al., 2000)	B16-FLT3L (RRID:CVCL_IJ12)
Experimental models: Organisms/Strains		
Mouse: C57BL/6J	Harty Laboratory and National Cancer Institute	#556

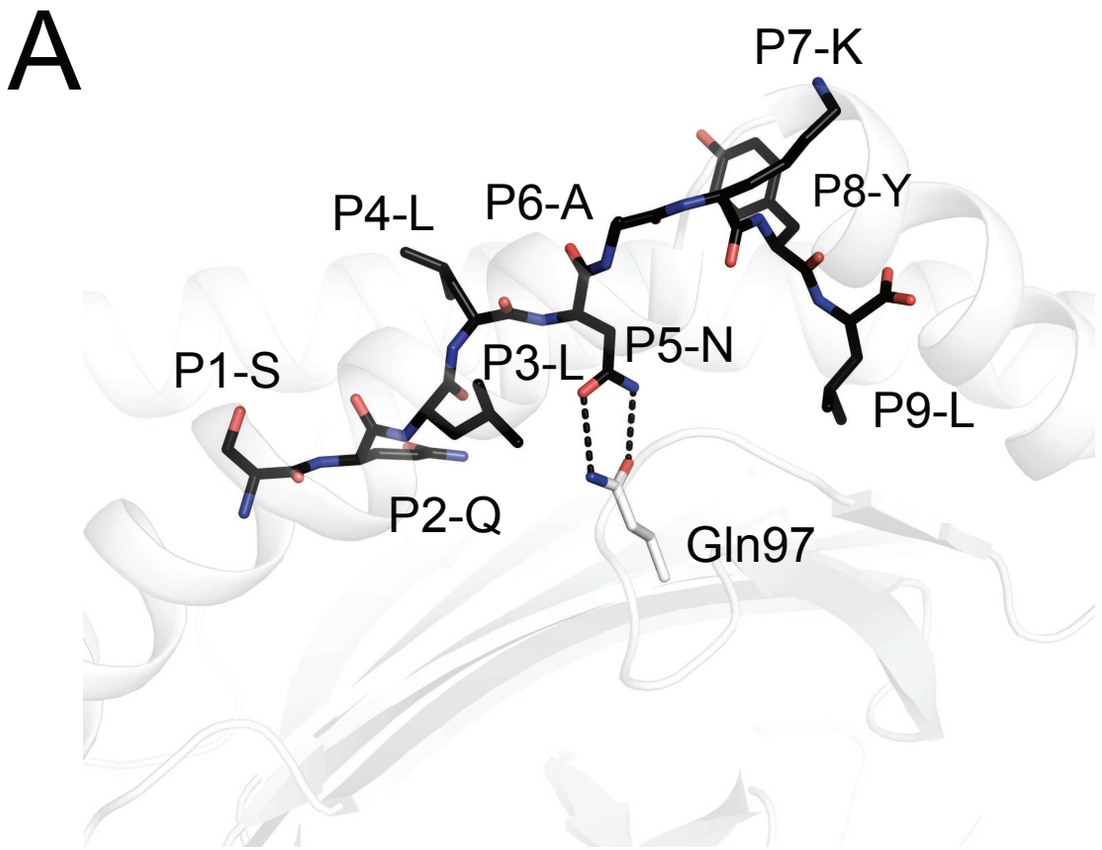
Deleted: PDB

Deleted: PDB

Mouse: CB6F1	National Cancer Institute	#566
Mouse: BALB/c	Harty Laboratory and National Cancer Institute	#555
Mouse: BALB.b	The Jackson Laboratory	001952
Mouse: (C57BL/6J x C57L) F1	This manuscript	NA
Mouse: C57L	The Jackson Laboratory	000668
Mouse: C57Bl/6 with $GAP50_{40-48}$-specific TCRβ chain retrogenic	This manuscript	NA
Mouse: C57Bl/6 with $TRAP_{130-138}$-specific TCRβ chain retrogenic	This manuscript	NA
P. berghei ANKA GFP-parasitized *Anopheles stephensi*	NYU Langone Health Insectary Core & Parasite Culture	NA
Recombinant virulent *L. monocytogenes*-$GAP50_{40-48}$	This manuscript	NA
Recombinant *actA-inlB- L. monocytogenes*-$TRAP_{130-138}$	Ref: (Doll et al., 2016)	NA
Recombinant *actA-inlB- L. monocytogenes*-$S20_{318-326}$	Ref: (Doll et al., 2016)	NA
Recombinant *actA-inlB- L. monocytogenes*-$GAP50_{40-48}$	Ref: (Doll et al., 2016)	NA
Software and Algorithms		
Diva 8	BD Biosciences	http://www.bdbiosciences.com
Sequencher 5.2.3	Gene Codes Corporation	http://www.genecodes.com/sequencher

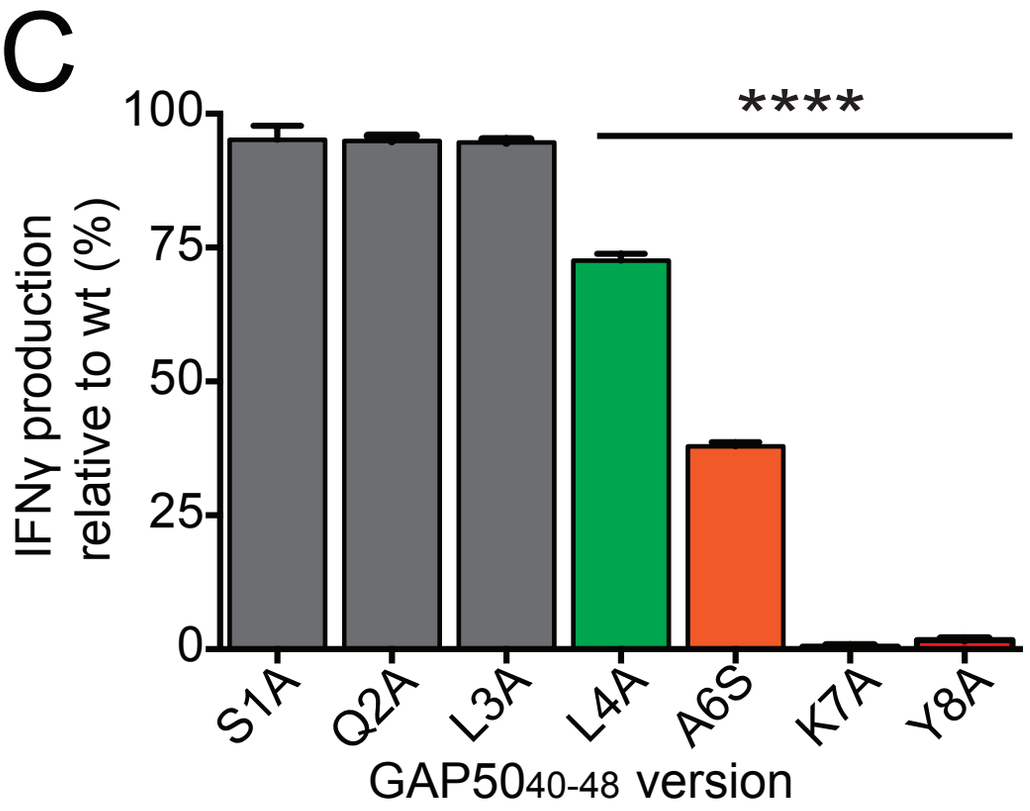
IMGT	*IMGT*®, the international ImMunoGeneTics information system	http://www.imgt.org
XDS	(Kabsch, 2010)	Available online http://xds.mpimf-heidelberg.mpg.de/
Scala	(Evans, 2006)	Available online http://www.ccp4.ac.uk/html/scala.html
CCP4	(Collaborative Computational Project, 1994)	Available online http://www.ccp4.ac.uk/
Phaser	(Read, 2001)	Available online http://www.ccp4.ac.uk/html/phaser.html
Coot	(Emsley and Cowtan, 2004)	Available online www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
Pymol	(DeLano, W.L., 2002)	Available online https://pymol.org/view
BIAevaluation	GE Healthcare	Version 3.1 https://www.biacore.com/lifesciences/

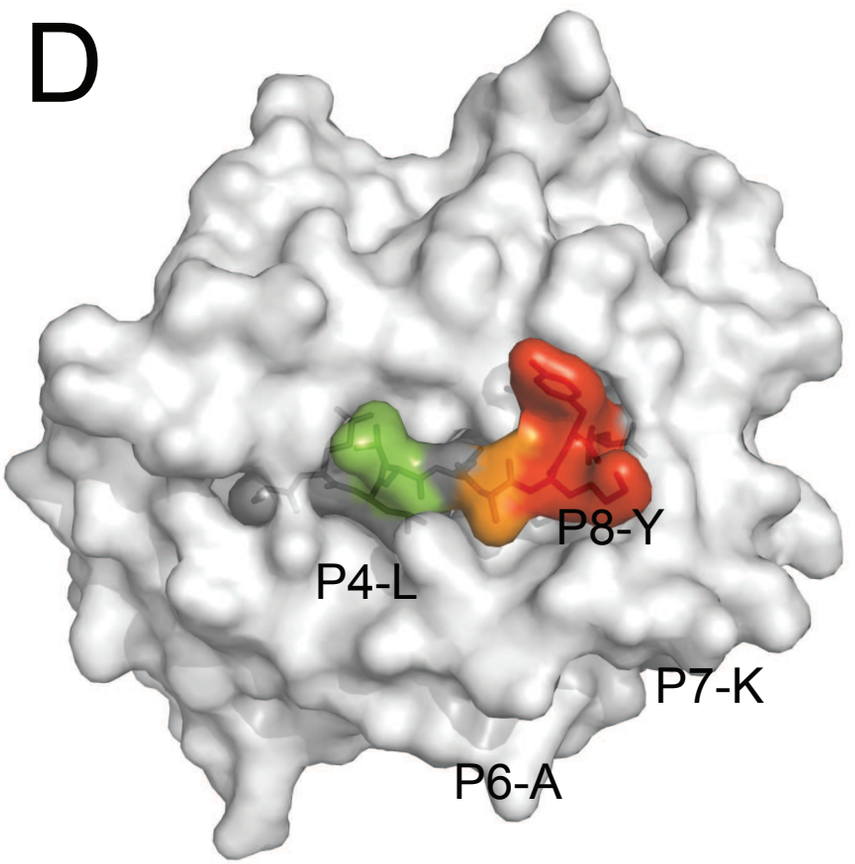

A
TRAP130-138
mouse #1
S20318-326
mouse #1
GAP5040-48
mouse #1
mouse #2
mouse #2
mouse #2
Vβ
2
3
4
5.1/5.2
6
7
8.1/8.2
8.3
9
10
11
12
13
14
17
B
GAP5040-48 tet.
CD8
99.5%
Vβ8.1/8.2
CD8
12.9%
10
Vβ8.1/8.2
CD8
C
Vβ8.1
D
GAP5040-48
mouse #1
Vα
12-3
9N-2
9D-4
7D-4
6-7/DV9
4D-4
6D-7
13D-4
6D-6
7-3
GAP5040-48
mouse #2
Vα
11
12-3
12D-3
8D-1
12-2
8-2
13-5
8D-2
1
10
19
9N-2
6D-6
8-1
14-1
GAP5040-48
mouse #3
Vα
9-1
4D-4
10
9D-4
9D-3
6-2
8D-2
15D-2/DV6D-2
5D-4
3-4
6D-7
8-1
E
V element
N region
D element
J element
CDR3β residue
C A S S D W G
% of total isolates
100
80
60
40
20
0
104 105 106 107 108 109 110 111 112 113 114
position

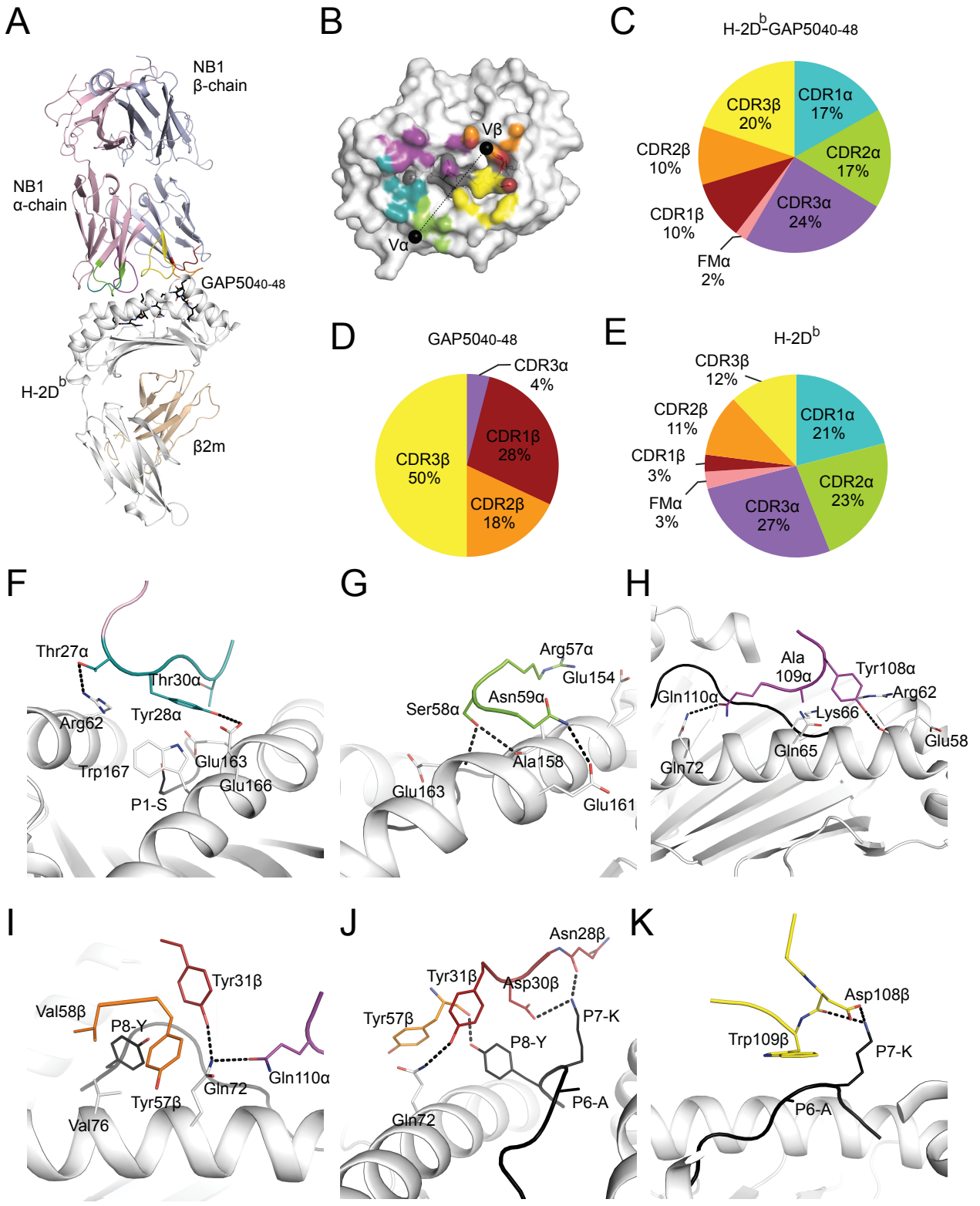


B

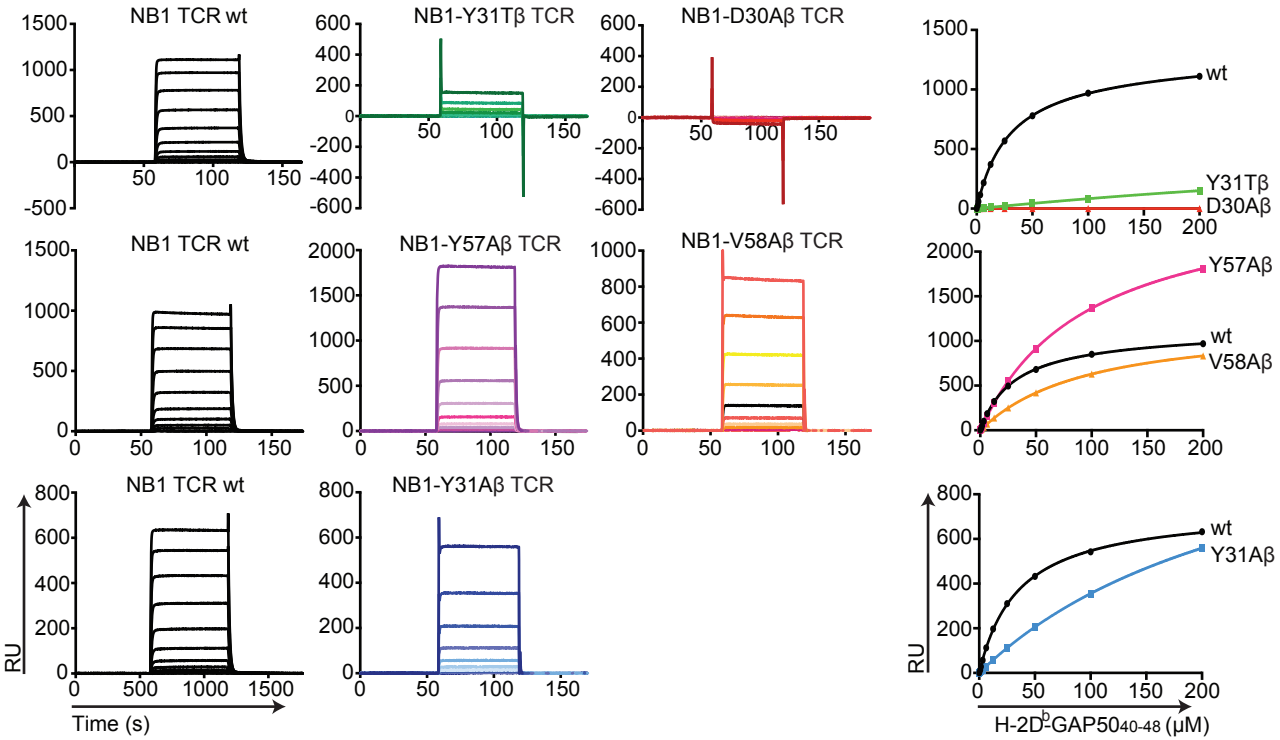
GAP50$_{40-48}$ version	amino acid sequence
wt	SQLLNAKYL
S1A	**A**QLLNAKYL
Q2A	S**A**LLNAKYL
L3A	SQ**A**LNAKYL
L4A	SQL**A**NAKYL
A6S	SQLLN**S**KYL
K7A	SQLLNA**A**YL
Y8A	SQLLNAK**A**L





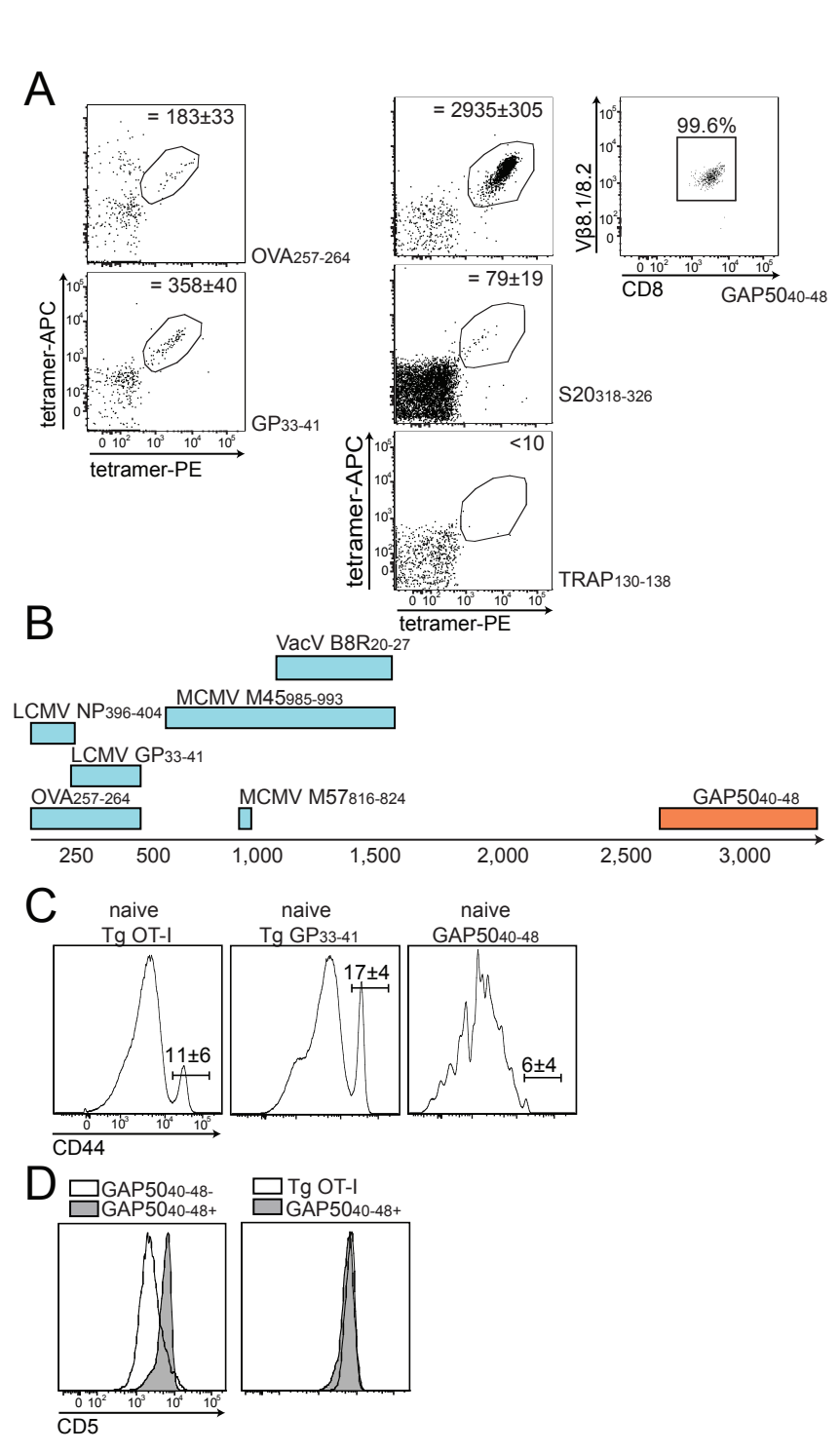

A
NB1
β-chain
NB1
α-chain
GAP5040-48
H-2Db
β2m
B
Vβ
Vα
C
H-2Db-GAP5040-48
CDR1α
17%
CDR2α
17%
CDR3α
24%
FMα
2%
CDR1β
10%
CDR2β
10%
CDR3β
20%
D
GAP5040-48
CDR3α
4%
CDR1β
28%
CDR2β
18%
CDR3β
50%
E
H-2Db
CDR1α
21%
CDR2α
23%
CDR3α
27%
FMα
3%
CDR1β
3%
CDR2β
11%
CDR3β
12%
F
Thr27α
Thr30α
Tyr28α
Arg62
Trp167
Glu163
Glu166
P1-S
G
Arg57α
Glu154
Asn59α
Ser58α
Ala158
Glu163
Glu161
H
Ala
109α
Tyr108α
Gln110α
Arg62
Lys66
Glu58
Gln65
Gln72
I
Tyr31β
Val58β
P8-Y
Gln110α
Gln72
Tyr57β
Val76
J
Asn28β
Tyr31β
Asp30β
Tyr57β
P7-K
P8-Y
P6-A
Gln72
K
Asp108β
Trp109β
P7-K
P6-A

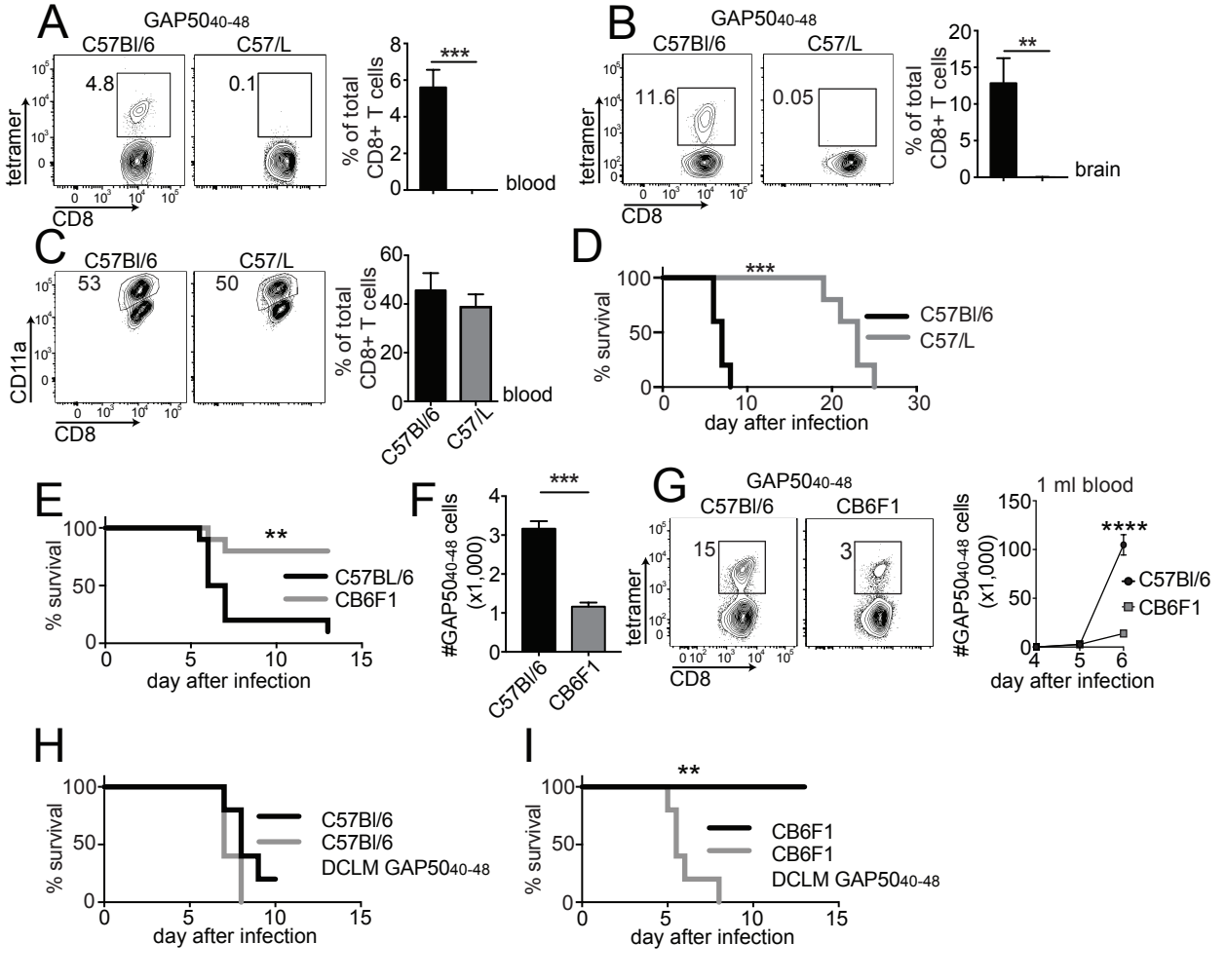
A

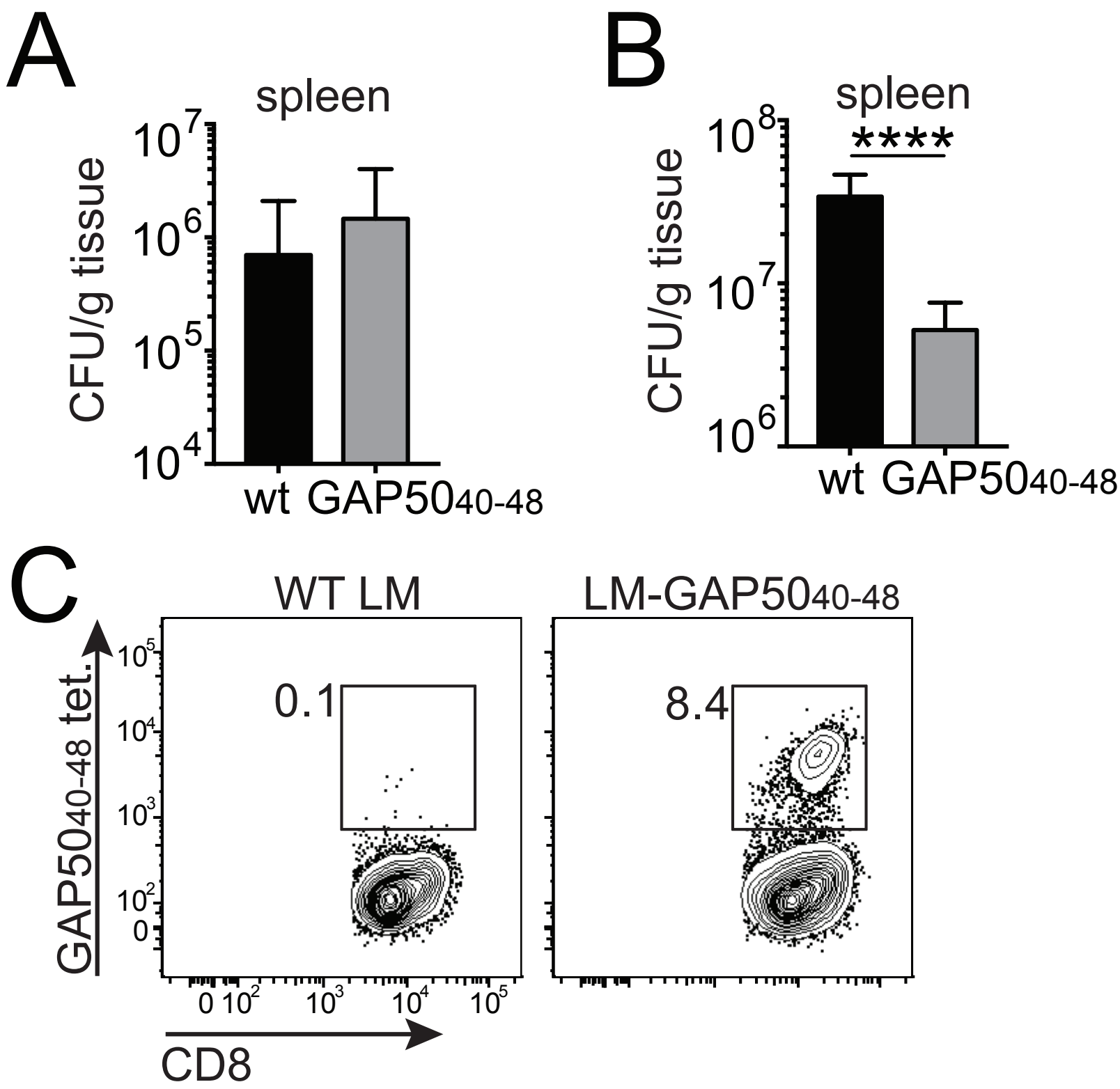


B

NB1 TCR	Kd$_{eq}$ (μM)
wt	19.0 ± 0.3
NB1 TCR mutants	**Kd$_{eq}$ (μM)**
Tyr31Alaβ	> 200
Tyr31Thrβ	> 200
Asp30Alaβ	NB
Tyr57Alaβ	59.0 ± 0.9
Val58Alaβ	59.7 ± 1.2


A
= 183±33
OVA257-264
= 358±40
GP33-41
tetramer-APC
tetramer-PE
= 2935±305
99.6%
Vβ8.1/8.2
CD8
GAP5040-48
= 79±19
S20318-326
<10
TRAP130-138
B
VacV B8R20-27
LCMV NP396-404
MCMV M45985-993
LCMV GP33-41
OVA257-264
MCMV M57816-824
GAP5040-48
250
500
1,000
1,500
2,000
2,500
3,000
C
naive
Tg OT-I
naive
Tg GP33-41
naive
GAP5040-48
11±6
17±4
6±4
CD44
D
GAP5040-48-
GAP5040-48+
Tg OT-I
GAP5040-48+
CD5


A
GAP50 40-48
C57Bl/6
C57/L
4.8
0.1
tetramer
CD8
% of total CD8+ T cells
***
blood
B
GAP50 40-48
C57Bl/6
C57/L
11.6
0.05
tetramer
CD8
% of total CD8+ T cells
**
brain
C
C57Bl/6
C57/L
53
50
CD11a
CD8
% of total CD8+ T cells
blood
C57Bl/6
C57/L
D
% survival
***
C57Bl/6
C57/L
day after infection
E
% survival
**
C57BL/6
CB6F1
day after infection
F
#GAP50 40-48 cells (x1,000)
***
C57Bl/6
CB6F1
G
GAP50 40-48
C57Bl/6
CB6F1
15
3
tetramer
CD8
1 ml blood
#GAP50 40-48 cells (x1,000)
****
C57Bl/6
CB6F1
day after infection
H
% survival
C57Bl/6
C57Bl/6
DCLM GAP50 40-48
day after infection
I
**
% survival
CB6F1
CB6F1
DCLM GAP50 40-48
day after infection


A
spleen
CFU/g tissue
10^7
10^6
10^5
10^4
wt GAP50$_{40-48}$
B
spleen
****
CFU/g tissue
10^8
10^7
10^6
wt GAP50$_{40-48}$
C
WT LM
LM-GAP50$_{40-48}$
GAP50$_{40-48}$ tet.
0.1
8.4
10^5
10^4
10^3
10^2
0
0 10^2 10^3 10^4 10^5
CD8

SUPPLEMENTAL MATERIAL

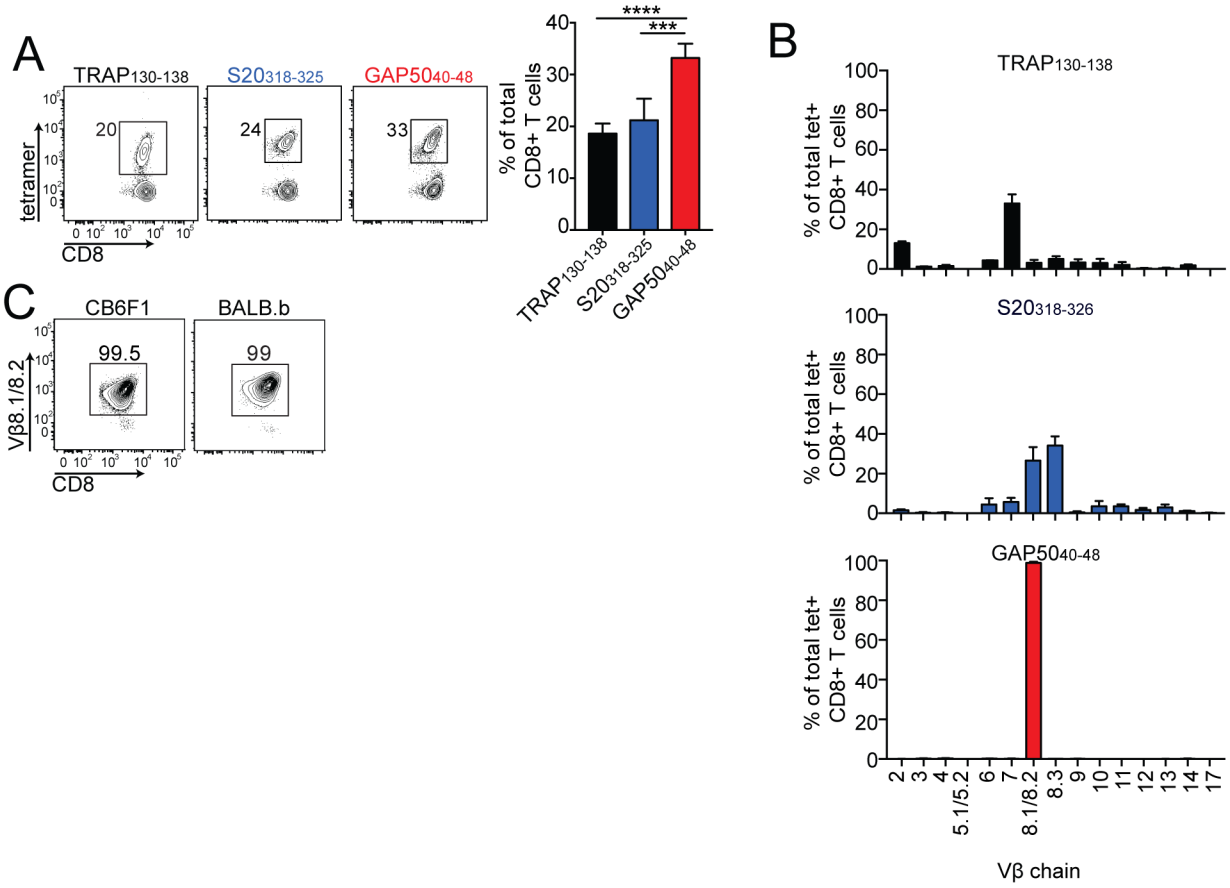


Figure S1 related to Figure 1: TCR Vβ screening of CD8$^+$ T cell populations specific for TRAP$_{130\text{-}138}$, S20$_{318\text{-}326}$ and GAP50$_{40\text{-}48}$

(A-B) Large CD8$^+$ T cell responses specific for TRAP$_{130\text{-}138}$, S20$_{318\text{-}326}$ and GAP50$_{40\text{-}48}$ were generated in C57Bl/6 mice using a DC-LM prime-boost strategy. (A) The magnitude of each antigen-specific response was assessed at day 45 post-infection in the blood of immunized mice using the corresponding MHC I tetramer (representative flow cytometry plots – left; cumulative graphs – right). Data represent four independent experiments (n = 5 mice/group). Bar graphs depict mean ± SD. Significance was assessed using a one-way ANOVA with post-hoc Tukey's multiple comparison test (***p<0.001, ****p<0.0001). (B) Distribution of various Vβ chains within different antigen-specific TCR repertoires, assessed in the blood of DC-LM-immunized mice and expressed as % of total tetramer$^+$ CD8$^+$ T cells. Data represent three independent experiments (n = 4–5 mice/group). Bars depict mean ± SD. (C) GAP50$_{40\text{-}48}$-specific

$CD8^+$ T cell responses were generated in CB6F1 and BALB/b mice using a DC-LM prime-boost strategy. Representative flow cytometry plots show Vβ8.1/8.2 antibody staining of $tetramer^+$ cells in the blood of immunized mice. Data represent two independent experiments.

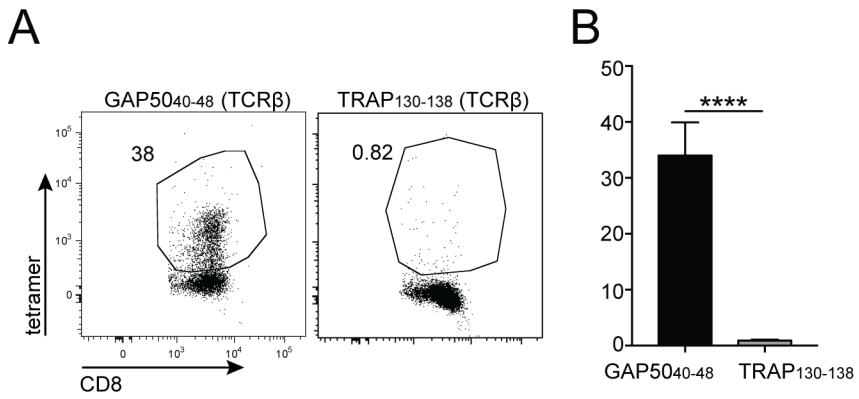


Figure S2 related to Figure 1: Tetramer staining of retrogenic CD8$^+$ T cells

(A, B) $GAP50_{40\text{-}48}$ and $TRAP_{130\text{-}138}$ TCRβ retrogenic mice were generated as described in the Methods. The success of TCR reconstitution was assessed in the blood of immunologically intact mice using specific MHC I tetramers (n = 5–6 mice/group). Representative flow cytometry plots (A) and cumulative graphs (B) showing tetramer staining of selected Vβ$^+$ CD8$^+$ T cells (Vβ8.1 for $GAP50_{40\text{-}48}$, Vβ9 for $TRAP_{130\text{-}138}$). Bars depict mean ± SD. Significance was assessed using a non-paired, two-tailed t test (****$p<0.0001$).

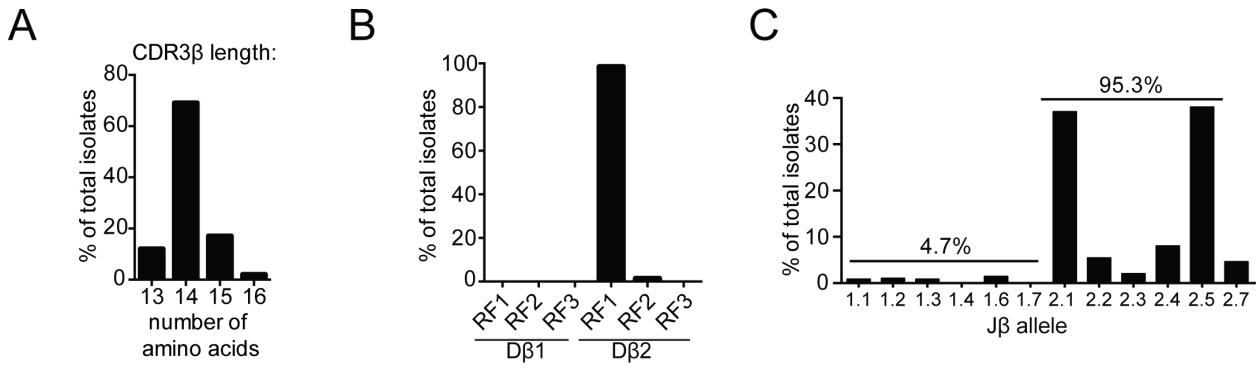


Figure S3 related to Figure 1: Conserved features of the GAP50$_{40-48}$-specific CDR3 β loop

(A–D) Genetic material used for sequence analysis was derived from viable CD3$^+$ CD4$^-$ CD8$^+$ GAP50$_{40-48}$ tetramer$^+$ cells sorted from DC-LM-immunized mice. (A) Length distribution of CDR3β sequences expressed as % of total number of molecular Vβ8.1$^+$ clones (n = 5 mice). (B, C) Usage of Dβ and Jβ elements in CDR3β sequences expressed as % of total number of molecular Vβ8.1$^+$ clones (n = 5 mice).

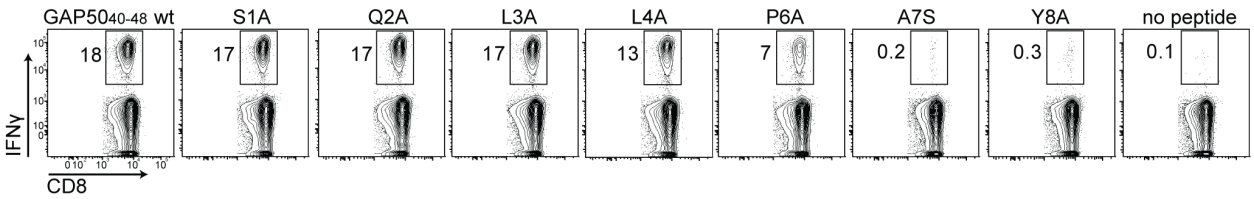


Figure S4 related to Figure 2: Identification of GAP50$_{40-48}$ residues critical for TCR activation

Representative flow cytometry plots showing IFNγ staining of GAP50$_{40-48}$-specific CD8^{+} T cells isolated from DC-LM-immunized mice and subsequently restimulated *in vitro* with *wt* or mutant GAP50$_{40-48}$ peptides.

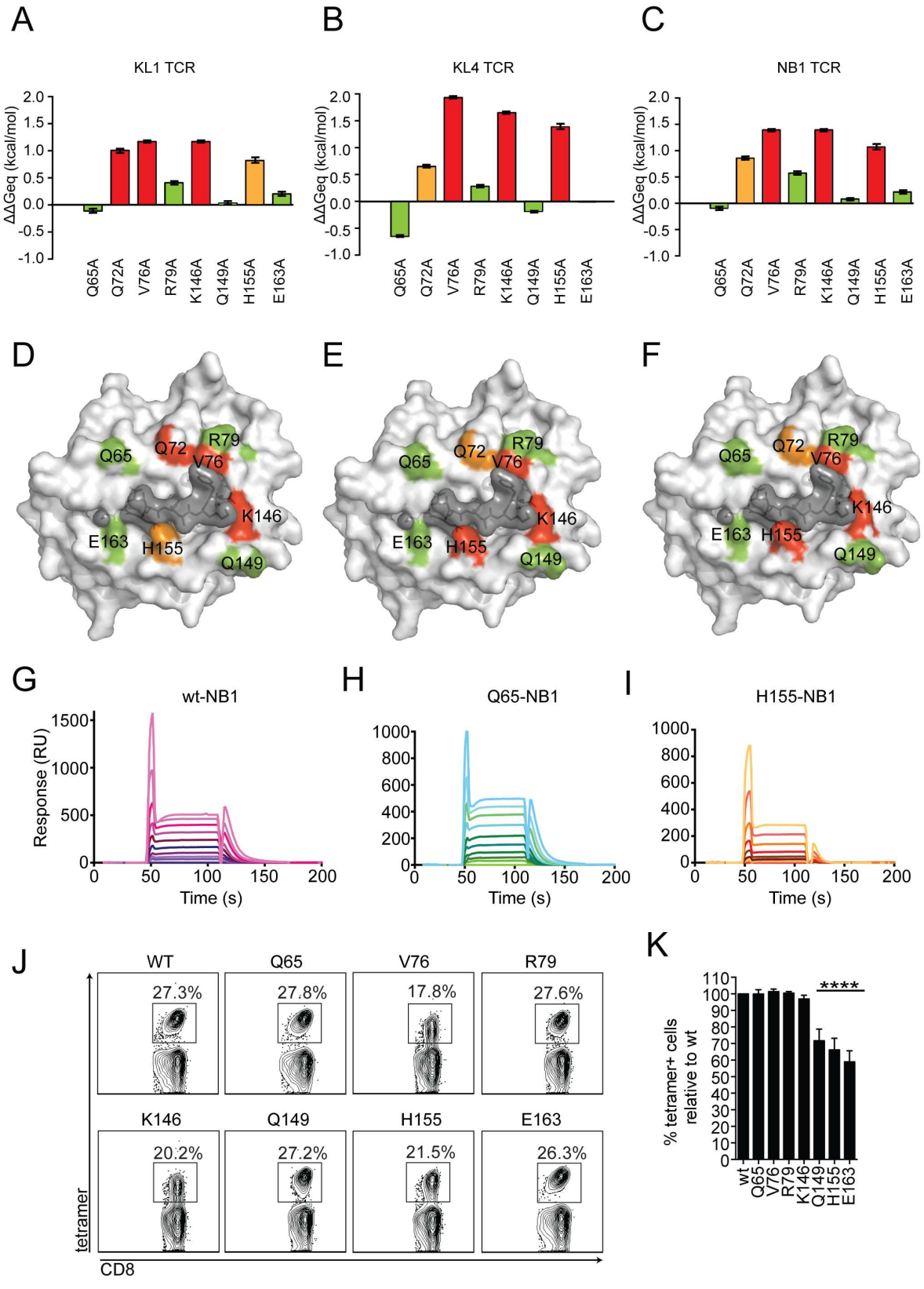

A
KL1 TCR
ΔΔGeq (kcal/mol)
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
Q65A
Q72A
V76A
R79A
K146A
Q149A
H155A
E163A
B
KL4 TCR
C
NB1 TCR
D
Q65
Q72
R79
V76
K146
E163
H155
Q149
E
F
G
wt-NB1
Response (RU)
1500
1000
500
0
0
50
100
150
200
Time (s)
H
Q65-NB1
1000
800
600
400
200
I
H155-NB1
J
WT
Q65
V76
R79
27.3%
27.8%
17.8%
27.6%
K146
Q149
H155
E163
20.2%
27.2%
21.5%
26.3%
tetramer
CD8
K
% tetramer+ cells relative to wt
****
wt
Q65
V76
R79
K146
Q149
H155
E163

Figure S5 related to Figure 3: Docking strategy of GAP50$_{40-48}$-specific TCRs

(A–I) SPR analysis of GAP50$_{40-48}$-specific TCRs isolated from the KL1, KL2 and NB1 CD8$^+$ T cell clones (Supplementary Table 5) to the indicated mutant H-2D^b molecules presenting the GAP50$_{40-48}$ peptide. Ala mutations were introduced at selected sites based on the crystal structure of the NB1 TCR-H-2D^b-GAP50$_{40-48}$ complex (Gln65, Val72, Gln72, Arg79, Lys146, His155 and Glu163). Ala mutation of Gln149 was used as a negative control. Ala substitutions were divided into 3 categories as follows: no impact (< 3-fold decrease in affinity) on TCR binding (green); moderate impact (3–5-fold decrease in affinity) on TCR binding (orange); and critical impact (> 5-fold decrease in affinity) on TCR binding (red). SPR results for H-2D^b Ala mutants loaded with the *wt* GAP50$_{40-48}$ peptide are shown for the KL1 TCR (A), the KL4 TCR (B) and the NB1 TCR (C). Experiments were conducted in duplicate. Bars depict SD. $\Delta\Delta G_{eq}$ = RT.ln (K_{deq}mut/K_{deq}wt), where mut designates mutant H-2D^b. (D–F) The surface of H-2D^b is shown in white, and the GAP50$_{40-48}$ peptide is shown in gray. (G–I) Representative SPR binding curves for the NB1 TCR with *wt* H-2D^b-GAP50$_{40-48}$ (G), Q65A H-2D^b-GAP50$_{40-48}$ (H) and H155A H-2D^b- GAP50$_{40-48}$ (I). The H-2D^b-GAP50$_{40-48}$ complexes were tested across a range of concentrations, indicated by the different colors, up to a maximum of 200 µM). (J–K) Splenocytes isolated from GAP50$_{40-48}$-immunized mice were stained with *wt* or mutant MHC I tetramers. Representative flow cytometry plots are shown in (J), and summary graphs are shown in (K). Data represent two independent experiments (n = 3 mice/group). Bars depict mean ± SD. Significance was assessed using a one-way ANOVA (****p<0.0001).

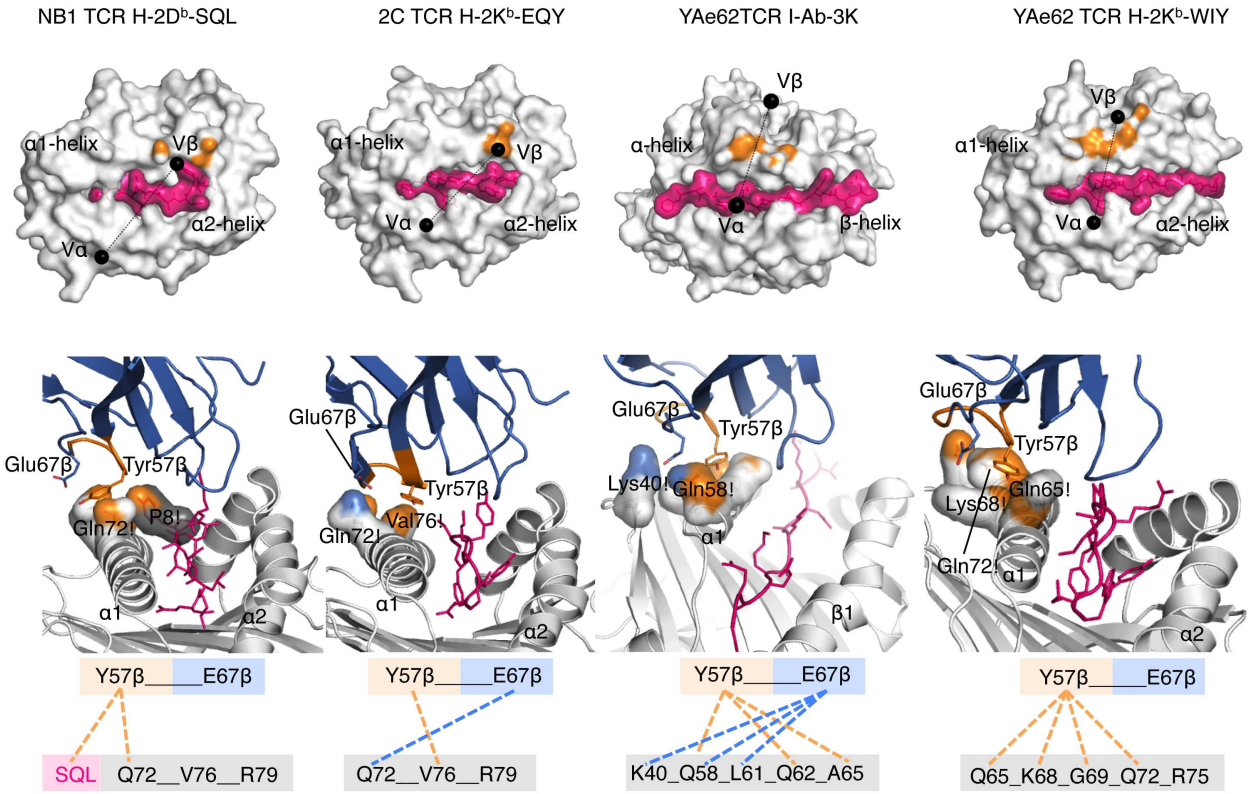


Figure S6 related to Figure 3: Interaction of the conserved Y57 and E67 of the Vβ8 genes with MHC class I and II molecules

From left to right: NB1 TCR-H-2D^{b}-SQL, 2C TCR-H-2K^{b}-EQY [PDB code: 2CKB, (Garcia et al., 1998)], YAe62 TCR-I-Ab-3K [PDB code: 3C60, (Dai et al., 2008)] and YAe62 TCR-H-2K^{b}-WIY [PDB code: 3RGV, (Yin et al., 2011)] complexes. The Top panels represent the footprint of the CDR2β loop for each TCR (orange) on the surface of their respective peptide (pink) and MHC (white) surface.

The middle panels represent a close up view of the interactions between the CDR2β loop (orange) and FWβ (blue) with their peptide (pink sitck) and MHC (white cartoon). The Tyr57β and the GLu67β are represented in stick, as well as the residues from the pMHC that are in contact with them.

The bottom panels show a representation of the interaction of the Tyr57β (orange dashes) and the Glu67β (blue dashes) with the residues from the peptide (pink box) and/or the MHC molecules (gray box).

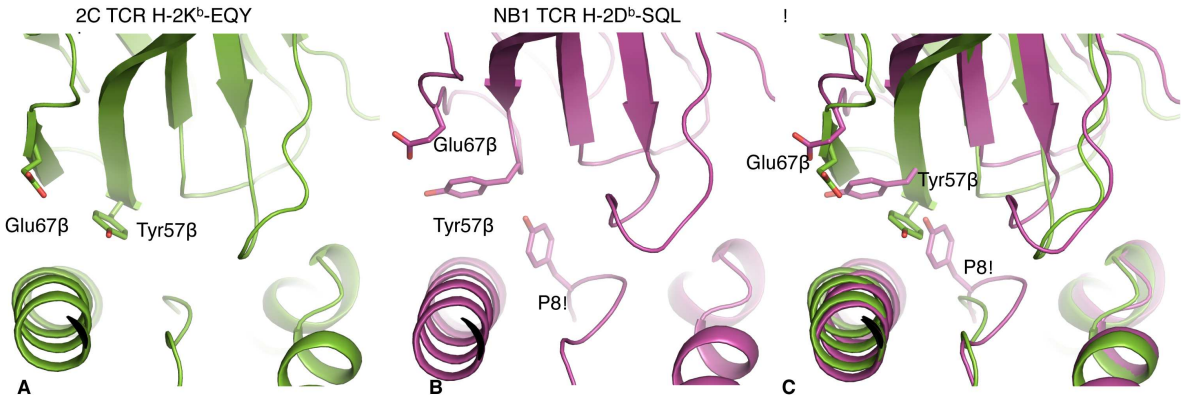


Figure S7 related to Figure 3: Different location of the Tyr57β between Vβ8.2+ 2C TCR and the Vβ8.1+ NB1 TCR

The panels show a close up view of the TCR Tyr57β and Glu67β interaction with the pMHC from the (A) 2C TCR-H-2K^{b}-EQY [PDB code: 2CKB, (Garcia et al., 1998)] complex (green), (B) the NB1 TCR-H-2D^{b}-SQL complex (purple), and (C) a superimposition of both of them. The C panel shows that the P8-Tyr of the SQL (purple stick) would clash with the location of the Tyr57β (green stick) of the 2C TCR.

Table S1 related to Figure 1: CDR3β sequences of Vβ8.1^{+} GAP50$_{40\text{-}48}$-specific transcripts

CDR3β (amino acid)	CDR3β (nucleotide)	Counts
CASSDWGTDAEQF	tgtgccagcagtgattgggggacagatgctgagcagttc	93
CASSDWGVQDTQY	tgtgccagcagtgactgggggggtccaagacacccagtac	21
CASSDWGVQDTQY	tgtgccagcagtgattgggggggtccaagacacccagtac	20
CASSDWGNQDTQY	tgtgccagcagtgattgggggaaccaagacacccagtac	16
CASSDWGVQDTQY	tgtgccagcagtgactgggggggtccaagacacccagtac	14
CASSDWGNYAEQF	tgtgccagcagtgactgggggaactatgctgagcagttc	11
CASSDWGVQDTQY	tgtgccagcagtgactgggggggtccaagacacccagtac	11
CASSDWGNQDTQY	tgtgccagcagtgactgggggaaccaagacacccagtac	8
CASSDWGGGQNTLY	tgtgccagcagtgactgggggggaggtcaaaacaccttg	8
CASSDWGNQDTQY	tgtgccagcagtgattggggtaaccaagacacccagtac	6
CASSDWGRDAEQF	tgtgccagcagtgactgggggcgagatgctgagcagttc	5
CASSHWGGDTQYF	tgtgccagcagccactggggtggggacacccagtacttt	5
CASSDWGNTGQLY	tgtgccagcagtgactggggaaacaccgggcagctctac	4
CASSDWGGVEQY	tgtgccagcagtgactgggggggcgttgaacagtacagt	4
CASSDWGASSYEQY	tgtgccagcagtgattgggggccagctcctatgaacag	3
CASSDWGVDAEQF	tgtgccagcagtgactgggggtagatgctgagcagttc	3
CASSDWGNYAEQF	tgtgccagcagtgattggggaaactatgctgagcagttc	3
CASSDWGGGQNTLY	tgtgccagcagtgattgggggggcggtcaaaacaccttg	3
CASSDFGTADTEVF	tgtgccagcagtgatttcgggacagcggacacagaagtc	2
CASSDWGAQDTQY	tgtgccagcagtgactggggagcccaagacacccagtac	2
CASSDWGGAYEQY	tgtgccagcagtgactgggggggcgcgtatgaacagtac	2
CASSDWGSGNTLY	tgtgccagcagtgattggggtctggaaatacgctctat	2
CASSDWGGQDTQYF	tgtgccagcagtgactgggggggacaagacacccagtac	2
CASSDWGGQNTLY	tgtgccagcagtgactgggggggtcaaaacaccttgtac	2
CASSDWGGYAEQFF	tgtgccagcagtgactgggggggctatgctgagcagttc	2
CASSDWGQDTQY	tgtgccagcagtgactggggcaagacacccagtacttt	2
CASSDWGTSPLY	tgtgccagcagtgactggggacctctccttgtacttt	2
CASSDWGNYAEQF	tgtgccagcagtgactggggaactatgctgagcagttc	2
CASSDWGGSQNTLY	tgtgccagcagtgattgggggggagtcaaaacaccttg	2
CASSDWGAAGNTLY	tgtgccagcagtgattgggggcggctggaaatacgctc	1
CASSDWGIYAEQFF	tgtgccagcagtgattggggatctatgctgagcagttc	1
CASSDWGQGSDYT	tgtgccagcagtgattggggacagggctccgactacacc	1
CASSDFQQDTQYF	tgtgccagcagtgatttccagcaagacacccagtacttt	1
CASSDWGGAETLY	tgtgccagcagtgactgggggggcgcagaaacgctgtat	1
CASSDWGDTGQLY	tgtgccagcagtgactgggggacaccgggcagctctac	1
CASSDWGGVEQY	tgtgccagcagtgactgggggggcgttgaacagtacagt	1
CASSDWGGQNTLY	tgtgccagcagtgactgggggggacaaaacaccttgtac	1
CASSDWGNYAEQF	tgtgccagcagtgactggggaactatgctgagcagttc	1
CASSDWGSAETLY	tgtgccagcagtgattgggggagtgcagaaacgctgtat	1
CASSDWGSQNTLY	tgtgccagcagtgattggggtagtcaaaacaccttgtac	1
CASSDWGVYAEQF	tgtgccagcagtgactgggggtctatgctgagcagttc	1

Blue: *V* gene; red: *D* gene; black: N insertion; purple, green, gray, orange, pale blue and pink: *J* genes.

Table S2 related to Figures 2 and 3: Data collection and refinement statistics

Data collection	**H-2D^b-GAP50$_{40-48}$**	**NB1 TCR-H-2D^b-GAP50$_{40-48}$**
Temperature (K)	100	100
Resolution range (Å)	48.56 - 2.20 (2.27 - 2.20)	38.82 - 2.1 (2.175 - 2.1)
Space group	$P2_1$	$P2_1$
Unit cell (Å) (°)	53.27, 159.92, 108.46 β= 92.33	115.74, 70.21, 116.49 β= 108.28
Total reflections	181134 (17680)	206330 (20547)
Unique reflections	91389 (9050)	103825 (10315)
Multiplicity	2.0 (2.0)	2.0 (2.0)
Data Completeness (%)	99.61 (97.23)	100 (100)
I/σ(I)	16.10 (4.31)	15.61 (2.42)
R_{pim}[a] (%)	3.9 (24.0)	3.4 (34.6)
Refinement statistics		
R_{factor}[b] (%)	19.4	19.3
R_{free}[b] (%)	24.1	23.5
Non-hydrogen atoms		
Protein	12641	12906
Water	1262	890
Rms deviation from ideality		
Bonds lengths (Å)	0.01	0.01
Bonds angles (°)	1.13	1.14
Ramachandran plot (%)		
Favored	99.3	95.0
Outliers	0	0.8

[a] $R_{pim} = \Sigma_{hkl} [1 / (N-1)]^{½} \Sigma_i | I_{hkl,i} - < I_{hkl} > | / \Sigma < I_{hkl} >$

[b] $R_{factor} = \Sigma_{hkl} | | Fo | - | Fc | | / \Sigma_{hkl} | Fo |$ for all data except 5% which were used for R_{free} calculation. Values in parentheses represent the highest resolution-shell.

Table S3 related to Figure 3 and S5: CDR3 sequences of GAP50$_{40-48}$ specific TCR clones

TCR	V α	CDR3 α	V β	CDR3 β
NB1	18	CATVYAQGLTF	8.1	CASS**DWG**DTGQLYF
KL1	17.1	CAMSPQGGSAKLIF	8.1	CASS**DWG**NQDTQYF
KL4	18	CATEGNNAGAKLTF	8.1	CASS**DWG**NQDTQYF

In IMGT nomenclature, Vα17.1 corresponds to TRAV16D/DV11, Vα18 corresponds to TRAV8-2, and Vβ8.1 corresponds to TRBV13-3.

Table S4 related to Figure 3: Contacts between the NB1 TCR and the H-2D^{b}-GAP50$_{40-48}$ complex

TCR region	TCR residues	H-2D^{b} residues	Bond type
CDR1α	Thr27-Oγ1	Arg62-NH2	VDW, HB
CDR1α	Tyr28-OH	Glu163, Glu166-Oε2, Trp167	VDW, HB
CDR1α	Thr30	Glu163	VDW
CDR2α	Arg57	Ala158	VDW
CDR2α	Ser58-Oγ	Ala158-O, Gly162, Glu163-N	VDW, HB
CDR2α	Asn59-Nδ2	Ala158, Glu161-Oε1, Gly162	VDW, HB
CDR3α-J	Tyr108-OH	Glu58-O, Arg62, Gln65	VDW, HB
CDR3α-J	Ala109	Arg62, Lys66	VDW
CDR3α-J	Gln110-Oε1	Gly69, Gln72-Nε2	VDW, HB
CDR1β	Tyr31-OH	Gln72-Nε2	VDW, HB
CDR2β	Tyr57	Gln72, Val76	VDW
CDR2β	Val58	Val76, Arg79	VDW
CDR3β-D	Asp108-Oδ2	Ser150-O	VDW, HB
CDR3β-D	Trp109	Ser150, Gly151, His155	VDW
CDR3β-J	Thr112-Oγ1	His155-Nε2	VDW, HB
TCR region	TCR residues	GAP50$_{40-48}$ peptide residues	Bond type
CDR3α-J	Ala109	Leu4	VDW
CDR3α-J	Gln110	Leu4	VDW
CDR1β	Asn28-O	Lys7-NZ	VDW, HB
CDR1β	Asp30-Oδ2	Lys7-NZ, Tyr8	VDW, SB
CDR1β	Tyr31	Tyr8	VDW
CDR2β	Tyr57-O	Tyr8-OH	VDW, HB
CDR3β-D	Asp108-Oδ1-O	Lys7-NZ	VDW, SB
CDR3β-D	Trp109	Ala6	VDW

CDR: complementary-determining regions; HB: hydrogen bonds (cut-off = 3.5 Å), VDW: van der Waals (cut-off = 4 Å), SB: salt bridge (cut-off = 5Å).

Table S5 related to Figures 3 and S5: Equilibrium binding affinities of different $GAP50_{40-48}$-specific TCRs and selected H-$2D^b$ Ala mutants loaded with *wt* peptide

MHCp complex	KL1 TCR (K_{deq} µM)	KL4 TCR (K_{deq} µM)	NB1 TCR (K_{deq} µM)
H-$2D^b$-$GAP50_{40-48}$	27.8 ± 0.3	7.5 ± 0.2	19.0 ± 0.3
H-$2D^b$-Q65A-$GAP50_{40-48}$	23.1 ± 1.0	2.5 ± 0.1	16.3 ± 0.3
H-$2D^b$-Q72A-$GAP50_{40-48}$	152.0 ± 4.0	22.4 ± 0.2	80.5 ± 1.3
H-$2D^b$-V76A-$GAP50_{40-48}$	> 200	196.0 ± 2.0	> 200
H-$2D^b$-R79A-$GAP50_{40-48}$	55.6 ± 1.0	12.0 ± 0.2	50.0 ± 1.4
H-$2D^b$-K146A-$GAP50_{40-48}$	> 200	121.5 ± 3.5	> 200
H-$2D^b$-Q149A-$GAP50_{40-48}$	29.3 ± 0.7	5.5 ± 0.1	21.8 ± 0.2
H-$2D^b$-H155A-$GAP50_{40-48}$	111.0 ± 11.0	77.7 ± 3.6	115.5 ± 11.5
H-$2D^b$-E163A-$GAP50_{40-48}$	39.2 ± 2.1	7.4 ± 0.1	27.4 ± 1.0

Data represent at least two independent experiments performed in duplicate using surface plasmon resonance (mean ± SEM).

Table S6A related to Figure 4: Mouse CDR1β sequence alignment

Mouse Vβ	CDR1β
8.1	NNH**DY**
10.1	LGH**D**T
28.1	MKH**D**S
13.1	SGH**D**T
9.1	MNH**D**T
6.1	FNH**D**T

CDR1 β residues of the Vβ8.1^{+} NB1 TCR critical for H-2D^{b}-GAP50$_{40\text{-}48}$ recognition are depicted in red (Asp30β and Tyr31β). Listed are mouse *V* β gene segments with a common Asp30β residue (in bold). Tyr31β is unique to the *V* β *8.1* gene.

Table S6B related to Figure 4: Mouse CDR2β sequence alignment

Mouse Vβ	CDR2β
8.1	S**YV**ADS
15.1	ST**V**NSAI
1.1	YS**V**KQL

CDR2β residues of the Vβ8.1^{+} NB1 TCR moderately important for H-2D^{b}-GAP50$_{40\text{-}48}$ recognition are depicted in blue (Tyr57β and Val58β). Listed are mouse *Vβ* gene segments with a common Val58β residues (in bold).